Geographic Information Technology Training Alliance (GITTA) presents:

Spatial Change Analysis

Responsible persons: Chloé Barboux, Claude Collet

Table Of Content

1. Spatial Change Analysis

The main objective of spatial change analysis is to describe the structure and the pattern distribution of spatial changes.
Most spatial change analysis methods are developed in image mode as this mode offers a simpler and structured description of spatial distributions, allowing the simultaneous analysis of:

- Change in spatial features (regions)
- Change in the spatial distribution of thematic properties (discrete or continuous)

Spatial change can be analysed at two levels:

- Through the spatial distribution analysis of thematic change indices
- Through the spatial dynamics modelling

There are various methods proposed to investigate the spatial dimension of time change. We have already presented some of them applied to the description of the spatial distribution of properties: for continuous (**B-AN, Lesson 3. Kontinuierliche Räumliche Variablen**) and discontinuous spatial distributions (**B-AN, Lesson 2. Discrete Spatial variables** and **I-AN, Lesson 2. Discrete Spatial Variables**). As presented in the introductory Unit of this Lesson, methods are not only dependant on the objective of the analysis or on the nature of the spatial distribution, but they are also related with the level of content of the thematic information. Table 3.1 summarises methods used to investigate spatial changes

SPATIAL DYNAMICS
1. For spatial distribution of change indices: Spatial filtering (Qual/Quant) Trend Surface Analysis (Quant) Double Fourier Series (Quant)
2. For spatial dynamics modelling: Cellular Automata (Qual) Spatial Diffusion Models (Quant)

Qual: qualitative data (nominal)
Quant: quantitative data (ordinal, cardinal)

Examples of spatial dynamics analysis methods

Learning Objectives

- You will be able to select appropriate methods and techniques for a specific context of spatial dynamics analysis.
- You will master the basic principles of various spatial dynamics methods and be able to access to related text references for in depth learning.

1.1. Spatial Distribution Analysis of Change Indices

How are change index values distributed throughout space? Is this distribution of change spatially organised? What is the pattern distribution of change properties?

The main objective of the spatial distribution of *change indices* [1] approach is to describe the structure and pattern of the change distribution throughout space.
We will be focusing on two techniques:
Spatial filtering and **Trend surface analysis**

1.1.1. Spatial filtering

The major objective of *spatial filtering* [2] process is to structure the distribution pattern into two components, a regional and a local one. *Low-pass filters* [3] are operators that *smooth* the spatial distribution of properties in order to enhance the regional component. Conversely the local component is extracted by retaining the local variations of properties with *high-pass filters* [4] that use *gradient operators*. As seen in the Section 3.3.1 of the Lesson 2 about spatial discrete distributions (**I-AN, Lesson 2. Discrete Spatial Variables**), the spatial context (the *neighbourhood* [5]) of any cell can be defined form three parameters:

- The *size* of the filtering window
- The *shape* of the filtering window
- The *proximity* to the central cell

The spatial filtering process can be applied in the following geographic information context:

- Information is in image mode (raster format)
- Properties are estimated for each cell in the image, either from an exhaustive description of space or through a regionalisation process
- Spatial distribution can be either continuous or discontinuous
- Properties can be either qualitative or quantitative

As the principles of spatial filtering methods are supposed to be known, let us illustrate the application of such a filtering process on images of change indices.

[1] A change index is an indicator derived from multitemporal measurements. It expresses the amount of change within a period of time. It can describe the change behavior of a set of features (global) or of individual features. It can result from a difference, a ratio, ...

[2] Procedures used to isolate spatial components: local and regional variations, trend. Typical filtering techniques make use of a moving window for the smoothing of local variations (low-pass filters) or their enhancement (high-pass filters)

[3] A technique to smooth the distribution of property values. In the spatial context it makes use of a moving window that defines the neighbourhood. A central tendancy value is computed within the moving window and assigned to the central cell (see High-pass filter)

[4] A technique to enhance the local distribution of property values. In the spatial context it makes use of a moving window that defines the neighbourhood. A gradient value is computed within the moving window and assigned to the central cell (see low-pass filter)

[5] The neighbourhood is an area surrounding a spatial feature or a spatial unit of observation (object or cell). Based on the spatial dependancy assumption, it stated that the neighbourhood of a feature influences its properties. A neighbourhood is defined by a range and a shape that encompass the feature up to a certain distance

First example: Dealing with a continuous spatial distribution, we want to analyse the spatial distribution of vegetation density change between 1984 and 2000 in the surrounding area of Fribourg City in Switzerland. A vegetation density index for the two dates was derived from Landsat TM remotely sensed images acquired during the month of July with a spatial resolution of 30m. The Normalised Difference Vegetation Index (NDVI) expresses the global density of vegetation within each pixel. Although this normalised ratio is aimed to compensate for external influences, its values cannot be considered as fully calibrated. However, the analyst wants to identify areas of significant changes in vegetation density between the two dates in order to interpret if it is either a temporary change due to different seasonal conditions or a permanent one. The animation below shows the process for identifying major areas of interest in this example.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

Second example: Dealing with a discontinuous spatial distribution, we want to analyse the spatial distribution of landcover types (qualitative level) during the period 1946-2001 in the area of Bulle in Switzerland. The geographical database was built by K. Al Ghamdi (Al-Ghamdi 2008) for his PhD study on the modelling of landcover change in this area. In order to identify landcover change between 1946 and 2001 we have derived a C index as a change indicator. C takes only two possible values: 0 for no change and 1 for any land cover change during this period of time. The animation below shows the process for the detection of all features corresponding to a change in landcover

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

1.1.2. Trend surface analysis

Principle

This approach is aimed to model the overall distribution of properties throughout space. It will sketch the global trend of distribution as a simplified surface.

Trend surface modelling [6] can be applied in the following geographic information context:

- Information can be either in image (raster format) or object (vector format) form
- Properties are estimated at sampled locations, described as a set of geographical locations
- Only spatially continuous distributions can be modelled
- Properties must be quantitative, measured at cardinal level

The principle of a trend surface model is a regression function that estimates the property value Pi at any location, based on the Xi,Yi coordinates of this location. The general function is:

[6] A regression function modelling the property values (Z) based on their location (X,Y) in space: Z = f(X,Y)

$$P_i = f(X_i, Y_i)$$

With:

P_i : property value at location i

X_i , Y_i : coordinate values at location i

f : regression function

A trend surface model is a particular case of a bivariate regression model with two independent variables, the coordinates X and Y and a dependent variable, the thematic variable P to be modelled. One can selected a *linear regression function* [7] (first order) or, if the spatial distribution is more complex, a *polynomial function* [8] (2nd, 3rd, …, or nth order). The modelled surface will correspond to respectively a flat oriented plane or a curved surface with an increasing number of curvatures.

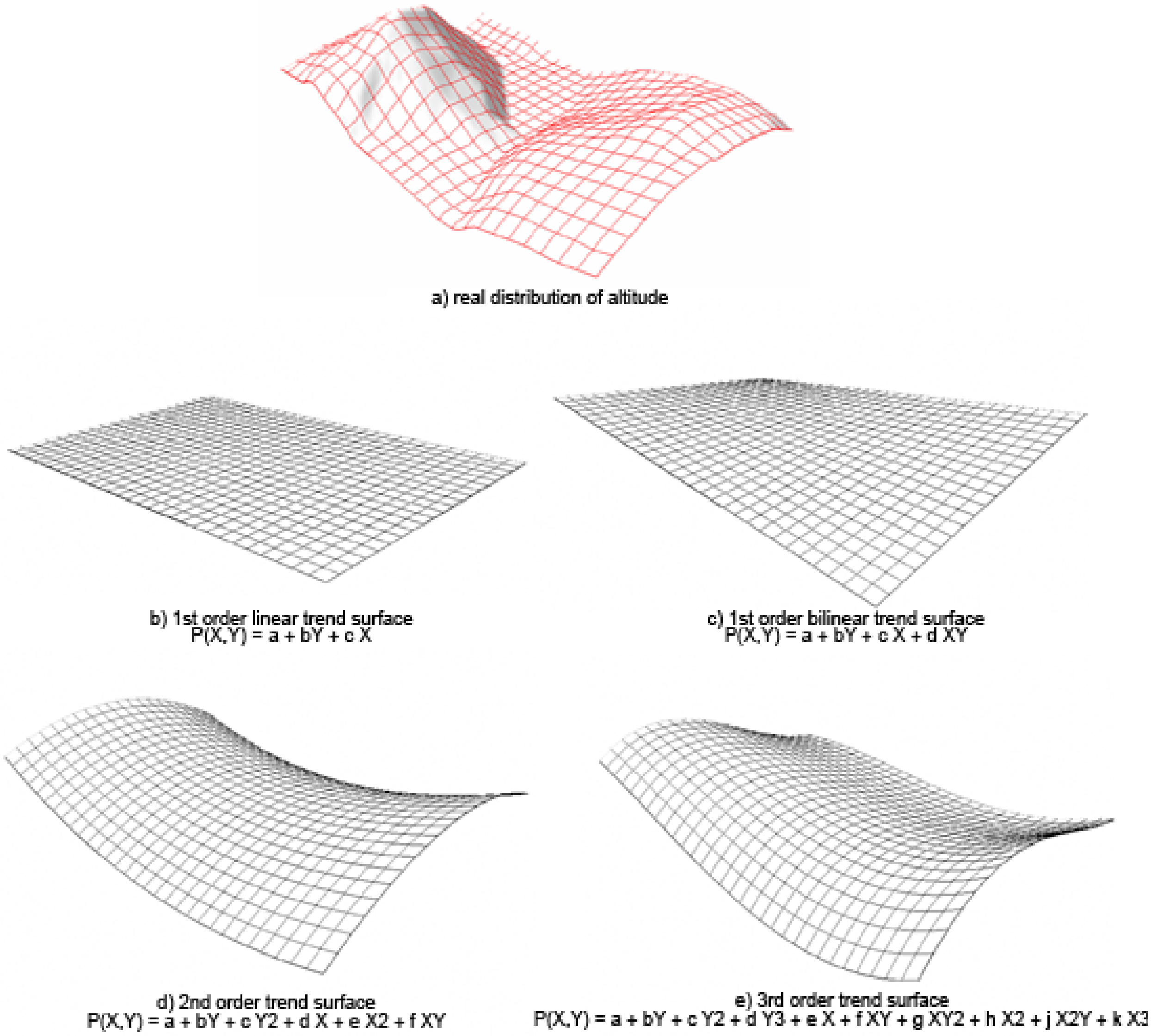
In order to illustrate the principles of trend surface modelling, let's take a phenomenon with a very obvious and observable spatial distribution: altitude. Let us suppose that we are starting our process of spatial distribution description with a sample of nn data point measurements, irregularly distributed throughout the study area to be described. One can identify 3 stages that are common to most modelling methods:

1. In the first step we select the most significant polynomial regression function that best explain the distribution of sample values. This is obtained by computing the F ratio that expresses the proportion of the total variance taken into account by the regression function. As this modelling approach is aimed to sketch the spatial distribution of properties with a continuous surface, it is recommended to limit the order of the regression function up to the fifth order. Such a surface requires 21 coefficients to be modelled. If the F ratio is not yet significant, this would mean that the distribution pattern of properties is simply too complex to be summarised with a surface function.
2. Once the regression model has been "calibrated" (i.e. estimation of function coefficient values and selection of the most appropriate function order), the regression function should be then applied to an independent set of sample points for validation purpose.
3. Finally, the selected regression function describes the considered trend surface that models the spatial distribution of properties. This function can then be used to estimate the property value at any location Xi,Yi within the study area.

The following figure illustrates the real distribution of altitude values within a study area as well as different trend surfaces modelling this distribution.

[7] A regression function that relates a dependant variable Y with one or several independant variables Xi in a linear manner. A first degree polynomial function is a linear function (see Polynomial regression function)

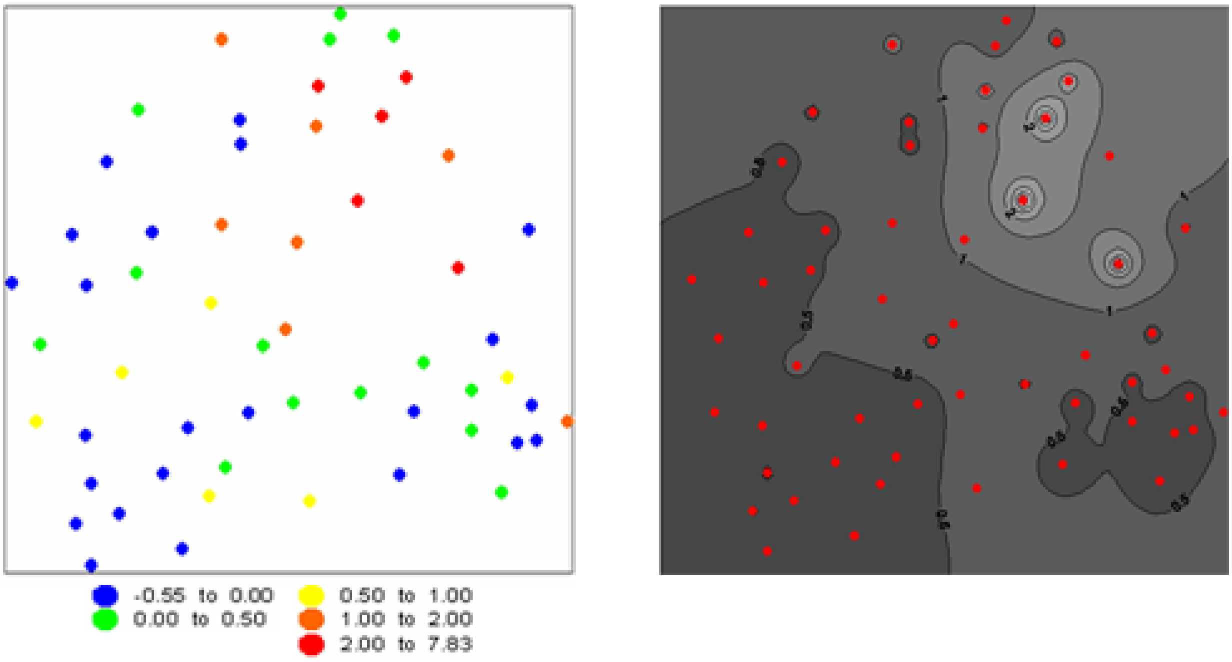
[8] A non-linear function that relates a dependant variable Y with one or several independant variables Xi. This function contains a combination of polynomes at different degrees or exponents (see Linear regression function)



Real spatial distribution of altitude in Exemplis study area and its modelled distribution based on trend surfaces of different orders

Illustration

Let us now illustrate the application of such a surface modelling process on a spatial sample of change index. Our objective is to summarise with the use of a trend surface the spatial distribution of change index values measured at different point locations in a study area. We have measured the population change of 55 localities in the district of the Sarine (in the Canton of Fribourg in Switzerland) during the period 1900-1986. The index change selected for the description of this evolution is the Normalised Difference (ND) as discussed in the section 2.1.1 of the Unit 2. Original X,Y coordinates were standardised in order to limit their unit of measurement and their influence on regression coefficient values.



a) Distribution of ND values

b) Interpolated surface from the 55 ND values

Spatial distribution of the normalised difference (ND) in population growth for the 55 localities in the district of the Sarine during the period 1900-1986.

Several regression functions were applied on this set of 55 points and Table summarises 3.2 their different parameters.

Regression function	Number of coefficients	%R^2	F ratio	Df numerator	Df denominator	Significance 95%
Order 1	3	16.3	5.07	2	52	Yes
Order 2	6	28.9	3.98	5	49	Yes
Order 3	10	39.7	3.29	9	45	Yes
Order 4	15	59.9	4.28	14	40	Yes
Order 5	21	73.7	4.77	20	34	Yes
Order 6	28	79.4	3.85	27	27	Yes
Order 7	36	91.2	5.60	35	19	Yes
Order 8	45	98.1	11.89	44	10	Yes

Summary of parameters produced to estimate the significance of each trend surface modelling function.

In order to test the significance of an individual trend surface model order, an *F* ratio should be derived from the coefficient of determination (%R^2) and the degrees of freedom *(Df)* associated with the fitted surface and its residuals (Davis 1986) (Unwin 1975). *F* ratio value is computed as follows:

$$F = (\%R2 / Dfnum) / ((100 - \%R2) / Dfden)$$

With:

%R2 : coefficient of determination of the regression function (fitted surface)

Dfnum : degrees of freedom associated with the regression function. They are equal to the number of coefficients in the function (ncoef) less one: Dfnum = ncoef - 1

Dfden : degrees of freedom associated with the residuals. They are equal to the number of observations (n) less degrees of freedom (Dfnum): Dfden = n - Dfnum

The F ratio values computed for each regression function order ranging from 1 to 8 are listed in the last table. They are all greater than their respective critical F value at 95 % confidence level and therefore express a significant trend. However, as our sample is made of 55 observations and as the number of coefficients increases rapidly with high function orders, it sounds reasonable to consider trend surface models up to the quintic order (order 5). From a statistical point of view there is a technique to assess the significance of contribution for selecting a higher order a lower one already significant. It is based on an F ratio value that expresses the extra contribution of an n+1 order over an n order (Davis 1986) (Unwin 1975). It is an interesting test as it compares the improvement of the fit with the increase in complexity of the fitting model. This F ratio value is computed as follows:

$$F = (\%R2 / Dfnum) / ((100 - \%R2) / Dfden)$$

With:

%R2 : coefficient of determination of the regression function (fitted surface)

Dfnum : degrees of freedom associated with the regression function. They are equal to the number of coefficients in the function (ncoef) less one: Dfnum = ncoef - 1

Dfden : degrees of freedom associated with the residuals. They are equal to the number of observations (n) less degrees of freedom (Dfnum): Dfden = n - Dfnum

The following illustrates the application of the test of increasing order significance for the 8 trend surface models applied to the 55 localities of the Sarine district.

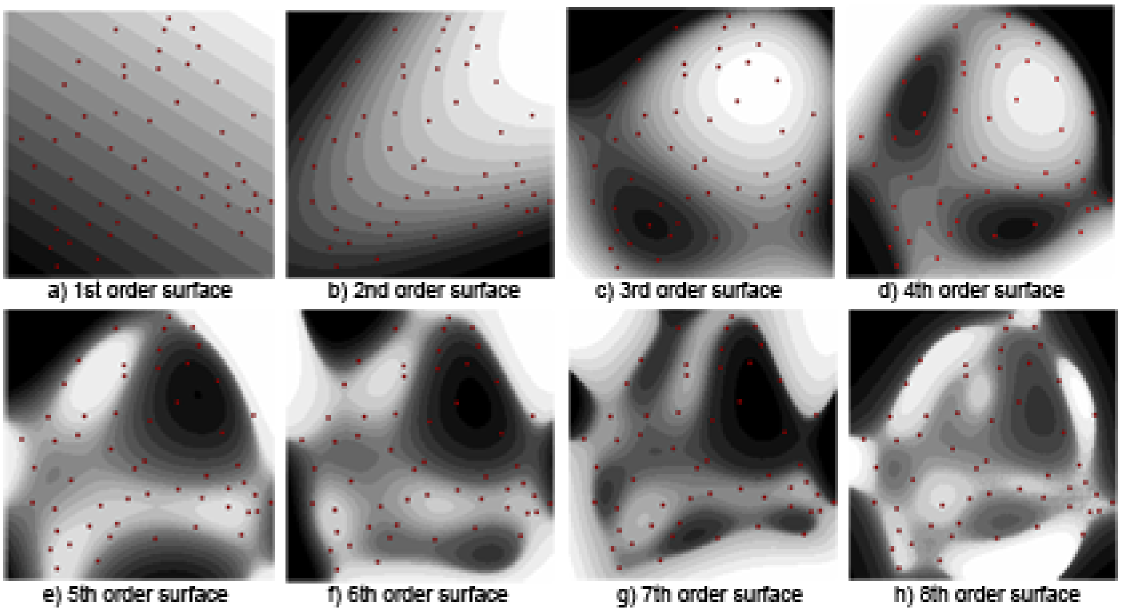
Order increase	Δ%R^2	Δ Coef	F numerator	F denominator	F ratio	Significance 95%
Order 1 to 2	12.6	3	4.20	1.45	2.89	Yes
Order 2 to 3	10.8	4	2.70	1.34	2.01	No
Order 3 to 4	20.2	5	4.04	1.00	4.03	Yes
Order 4 to 5	13.8	6	2.30	0.77	2.97	Yes
Order 5 to 6	5.7	7	0.81	0.76	1.07	No
Order 6 to 7	11.8	8	1.48	0.46	3.18	Yes
Order 7 to 8	6.9	9	0.77	0.19	4.04	Yes

Test of significance of the order n+1 against order n.

The interpretation of significance of trend functions listed before (**Table**) leads to the conclusion that the 8th order is the most suitable for modelling the spatial distribution as it takes into account 98.1% of the point distribution. This conclusion is confirmed by the test of significance about the additional contribution of the order 8 against the previous order (**Table**). However, when comparing the interpolated surface (**Figure**) with the modelled trend surfaces illustrated in the next **figure**, one can observe several strong discrepancies and artefacts generated by trend models, particularly with higher orders:

- Outside the sampled region estimated values from trend surfaces become unreal. This is known as edge effects. One should discard these regions by masking or obtain additional sample points. Such artefacts are observable from the 3rd order.
- Trend surfaces tend to fit better and better the interpolated distribution up to the 4th order. Then the distribution of values tends to be more and more different from the interpolated distribution. This discrepancy is mainly caused by the limited number of sample points compared with the rapid increase in number of coefficients attached to high order functions. Ultimately, when the number of coefficients reaches the number of observations, the trend surface will exactly fit all the observation values, but the remaining locations of this modelled surface will certainly be inconsistent.

Based on these general issues and keeping in mind that a relevant trend surface model is aimed to summarise the overall spatial distribution of properties, one should concludes that the most appropriate model for the description of this distribution could be the 4th order function. It explains around 60% of the overall variations and contains 15 coefficients, almost four times less than the number of observations.



Trend surfaces of different orders modelling the spatial distribution of population growth index during the period 1900-1986 in the district of the Sarine.

1.2. Spatial Dynamics Modelling

How to model spatial changes, movements and accessibility as the results of spatial processes?

Spatial dynamics modelling [9] is a main concern with the analysis of property changes process within space as well as movements and accessibility in space. This is a very wide and complex topic and task. The purpose of this section is to introduce you to the world of spatial modelling by highlighting some concepts, methodologies and current developments. It is of course impossible within a single teaching unit to cover in depth such a vast domain but certainly to propose a taste of it through introductory comments and illustrations.

1.2.1. Property changes in space

In section**Spatial distribution analysis of change indices** we were considering spatial features as permanent during the studied period and were focusing on the description of their property changes. In this section, we will now consider the study area as a whole and then model the spatial distribution of properties of a phenomenon and its change through the considered period of time. With this approach, spatial features are resulting from the spatial distribution of properties at each moment. In order to build up these spatial features, it is necessary to describe space with regular units of observation rather than existing spatial features, such as regular cells in image mode. The objective of this approach is to model the process of spatial property changes in order to describe past changes to the present time, but also to forecast future changes.
As previously discussed and illustrated, one should differentiate between phenomena with a continuous spatial distribution from those with a discontinuous (discrete) spatial distribution. This is important because of the differences in rules pertaining to spatial autocorrelation. Let us differentiate between two general approaches ruled by the nature of the spatial distribution of the considered phenomenon:

- *Discontinuous spatial distribution* [10] : As the number of possible properties is limited, they are mainly expressed at a nominal level with a categorical content. In this situation we are involved with the analysis of a change of state for each unit of observation. Each cell has a specific state at beginning of the process and the model has to evaluate the probabilities to change to another state according to probabilities or possibilities of occurrence as well as to the neighbourhood properties. In this situation the spatial dependency is limited to a local neighbour. This approach can be modelled with the use of cellular automata. **(link to lesson)**
- *Continuous spatial distribution* [11] : In this context the number of possible properties is very large or even infinite. The assumption is that the property at a specific location (a cell) is influenced by the proximity to other cells up to a threshold distance of influence. Throughout time the model should express a change of intensity (quantitative). Such situations assume a high spatial dependency (autocorrelation) and can be modelled by contagious spatial diffusion models **(link to lesson)**.

---

[9] Procedure used to summarise the changes in the spatial distribution of phenomenon properties throughout time (see Spatial dynamics)

[10] Properties of a phenomenon can be distributed throughout space in a continuous or discontinuous manner. A discontinuous (discrete) spatial distribution is characterised by strong changes in property values throughout space. Places with value changes delineate spatial feature boundaries (point, linear or areal) (see Continuous spatial distribution)

[11] Properties of a phenomenon can be distributed throughout space in a continuous or discontinuous manner. A continuous spatial distribution is characterised with gentle changes in property values throughout space. This is due to a continuous spatial dependancy throughout space. Thus the spatial distribution corresponds to a continuous surface rather than a set of spatial objects (see Discontinuous spatial distribution)

1.3. Spatial Dynamics - Discontinuous case

Cellular automata [12] were invented in the late 1940s by two mathematicians, John von Neumann and Stanislaw Ulam, working at the Los Alamos National Laboratory in the United States. Cellular Automata (CA) are dynamic systems which are discrete in space and time, operate on a uniform, regular lattice and are characterised by "local" interactions.
A CA system consists of a regular grid of cells; each can be in one of a finite number updated synchronously in discrete time steps:

- The state of a cell is determined by the previous states of the surrounding neighbourhood of cells.
- Each cell in a regular spatial lattice can have any one of a finite number of states.
- Local rule: the state of a cell at a given time depends only on its own state one time step previously and states of its nearby neighbours at the previous time step. The state of the entire lattice advances in discrete time steps

Different kind of problems can be approached using cellular structure and rules: *spatially complex systems* (e.g., landscape processes), *discrete entity modelling in space and time* (e.g., ecological systems, population dynamics) or *emergent phenomena* (e.g., evolution, earthquakes)
CA consist of different elements, they are:

- Cell Space: it is one cell that can be in any geometric shape.
- Cell State: can represent any spatial variable. For greater flexibility into CA, two groups of cell states integrated, fixed and functional.
- Time Steps: CA evolve at a sequence of discrete time steps
- Transition Rules: a transition rule specifies the state of cell before and after updating based on its neighbours conditions. In the classic CA, transition rules are deterministic and unchanged during evolution. Now the rules are modified into stochastic expressions and fuzzy logic controlled methods.
- Spatial neighbourhood: a *limited neighbourhood* [13] (Von Neumann) including 4 adjacent cells or an *extended neighbourhood* [14] (Moore) including the 8 adjacent cells

---

[12] Cellular automata are dynamic systems used to model spatial changes such as spatial diffusion, movements, … Cellular automata are dynamic agents obeying to rules that modify properties in space

[13] The neighbourhood is an area surrounding a spatial feature or a spatial unit of observation (object or cell). A limited neighbourhood generally corresponds to a set of spatial features (objects or cells) that are contiguous to the central feature in a limited sense: they at least share one side with it (see spatial neighbourhood and extended neighbourhood)

[14] The neighbourhood is an area surrounding a spatial feature or a spatial unit of observation (object or cell). An extended neighbourhood generally corresponds to a set of spatial features (objects or cells) that are contiguous to the central feature in a wider sense: they at least share one vertex with it (see spatial neighbourhood and limited neighbourhood)

The two considered spatial neighbourhoods: limited neighbourhood (Von Neumann) and extended (Moore).

As summarised by A.K. Singh (2003), a cellular automata model can be represented as the following quadruple:

(U, S, N, T)

With:

U: universe (cell space or lattice)
S: set of all possible states that a cell can reach
N: neighbourhood of a cell
T: a set of transition rules

Then the state of a cell at time (t+1) is a function of it state at time t of its neighbourhood and of the set of transition rules:

$$S_{t+1} = f(S_t, N, T)$$

With:

S_{t+1} : a state of a cell at time (t+1)
f : a function of a cell
St : a state of a cell at time (t)
N : the neighbourhood of a cell
T : a set of transition rules

1.3.1. Example of CA : Game of Life

Let us illustrate this with a very simple but famous CA model: the *Game of Life* [15]. It is a cellular automaton invented by Cambridge mathematician John Conway in the late 1960s. The neighbourhood is consisting of the nearest 8 cells to a cell on a two-dimensional grid of cells.
The space of the Game of Life is an infinite two-dimensional orthogonal grid of square cells, each of which is in one of two possible states, live or dead. Every cell interacts with its eight neighbours, which are the cells that are directly horizontally, vertically, or diagonally adjacent. At each step in time, the following transition rules take place:

- Any live cell with fewer than two live neighbours dies, as if by loneliness.
- Any live cell with more than three live neighbours dies, as if by overcrowding.

[15] A famous computer simulation illustrating the birth, growth and death of cells distributed throughout a regular gridded space. It demonstrates the use of a cellular automata model, using very simple rules, to simulate a spatial diffusion process

- Any live cell with two or three live neighbours lives, unchanged, to the next generation.

Many types of different patterns occur in the Game of Life, some of them are static patterns ("still lives"), repeating patterns ("oscillators"), and patterns which translate them self across the board ("spaceships").
We can break the patterns in three categories:

1. the block and boat are still lives
2. the blinker and toad are 2-phased oscillators , while pulsar is the most common period 3 oscillator
3. glider and lightweight spaceship (LWSS) are spaceships that move across the grid as time goes on

EXERCISE:

This is a PC based Game of Life program. Just click on the link below and you will be redirected to the program, you can choose from many patterns from the list in the lower left of the window and run the game and notice how these patterns move and take shapes, some of them keep on going either moving over the grid or moving in the same space over time. Some freeze after some time and some just disappear. You can create your own shapes and see how many generations they can last or change or even die, use the transition rules for the Game of Life to create new patterns so that you can understand first hand the behaviour of evolving shapes from time to time. As you will notice, some shapes live forever and some short or long life cycles. **The Game of Life**

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

1.3.2. Example of CA using Markov Chain

To illustrate the technique of modelling landcover change by CA, we will present a relatively simple but effective model that combines cellular automata spatial rules with Markov chain transition rules. For example, this approach is proposed as a module called MARKOV / CA_MARKOV in the raster GIS IDRISI (Eastman 2008). The procedure models predictive *Land Use Cover (LUC) changes* [16] within two stages

Stage 1

A *Markov chain analysis (MARKOV)* [17] is performed in order to estimate the transition matrix between the two past and documented dates (date 1 and date 2) and to estimate probabilities of change for the third date (date 3) to be predicted. Input and output parameters for this analysis are the following:

- The LUC distribution for the two dates is provided as two images

[16] The cover or the use of the earth surface is described with a set of category types. Throughout time these cover or use categories can change. The objective of a LUCC analysis is to describe, to understand or to predict such changes in the spatial distribution of landcover or landuse

[17] A technique to estimate the probability of occurrence from any original state to any final state after a specific sequence of n time steps. It makes use of transition matrices

- The interval of time between the two documented dates (date 1 and date 2) as well as the one between the second date and the date to be predicted (date 2 and date 3) are expressed as regular time steps (iterations)
- A mask image can be introduced in order to limit the development and change to another LUC category due to constraint rules. This will modify the transition probability matrix values
- A transition probability matrix is produced. It expresses the possibility that a cell of a given LUC category will change into any other category
- A transition area matrix is derived. It contains the total area (in cells) expected to change in the next time period
- A group of conditional probability images are generated, one image for each category. They express the probability that each cell will belong to the designated category in the next time period.

Stage 2

A Cellular automata predicting model (CA_MARKOV) estimates the spatial distribution of landcover at a later date (date 3):

- Using the output data produced by the Markov chain analysis, the predicting model will apply a contiguity filter to “grow out” landcover from date 2 to a later time (date 3).
- This CA filter develops a spatially explicit contiguity-weighting factor to change the state of a cell based on its neighbours.

The data used in this illustration are landcover map layers of 1952, 1974 and 1993 for the town of Bulle in the canton of Fribourg in Switzerland. As already previously commented, this dataset has been developed by K. Al-Ghamdi in the context on his PhD research work on LUCC modelling (Al-Ghamdi 2008). The objective of this illustration was to analyse the LUC change during the period from 1952 to 1974 and based on this modelling to predict the evolution for the date 1993. The ground truth 1993 LUC layer will then be used to assess the performance of the procedure. As already illustrated in figure 3.4 for the dates 1946 and 2001, six categories of landcover are identified: lake/pond, river, wetland/marshland, forest, agricultural/open field and urban. Linear features, such as the road network, were not included into the analysis as the 10m resolution of the LUC would overestimate its impact and development.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

Several comments can be made about this spatial modelling of LUCC:

- The mask image limits the development of the land cover during the period 1974-1993 to only two categories: urban and agriculture. Remaining categories are considered as protected from the land planning politics. Therefore, the land competition is limited to urban and agricultural development.

- From last slideshow one can observe that urban growth mainly occurs at the expense of agricultural lands and open fields. The expansion of urban areas is modelled on the edges of already urbanised areas, due to the contiguity constraint assigned to the model rules. This type of growth is called “organic growth”.
- f) shows discrepancies between the modelled and the real landcover distribution in 1993. The lack of urban expansion corresponds to new urbanised areas (in blue) that are compensated by extra contiguous zones as part of the organic growth. The inability to model new seeds of development is one of the strong limitations of such a model.

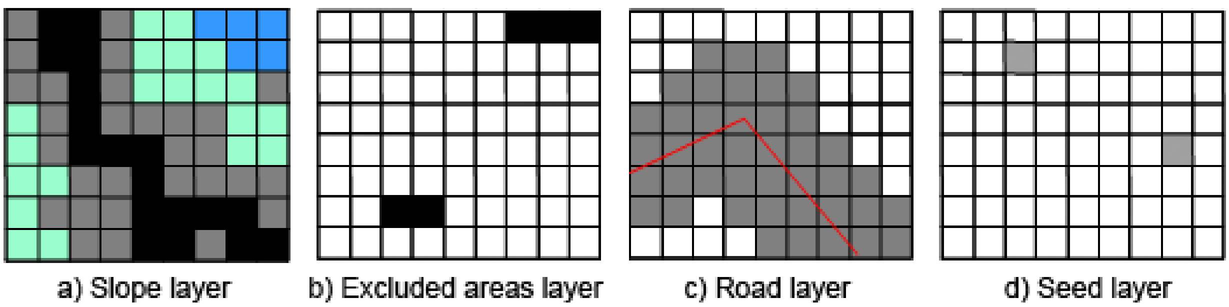
Among other limitations one should identify the inability to take into account the influence of the road network that could potentially generate new “seeds” of development.

1.3.3. Example of CA : SLEUTH, a more complex one

Now we shortly present a CA model proposed by Clarke et al. (1998). This CA simulation model was developed to predict urban growth in the San Francisco Bay area. They have used multiple data sources: a topography map, historical maps of highway development from 1920 to 1978, existing settlement distributions and their modification over time.
The SLEUTH name is the acronym for Slope, Landuse, Exclusion, Urban, Transportation and Hillshading. These are factors controlling the urban growth process.
The control parameters of the model are allowed to self-modify, where the CA adapts to the circumstances it generates, especially, during periods of rapid growth or stagnation. The model accumulated probabilistic estimates based on Monte Carlo methods.

The next figure illustrates the four input data layers used by *SLEUTH model* [18] for simulating urban growth:

- *Slope layer:* Slope value for every cell location is derived from a DEM. It is used to determine the slope-resistance weighting.
- *Excluded areas layer:* Cells entirely exempt from the growth process, including oceans, lakes, and protected areas such as national parks and wetlands.
- *Road layer:* A binary array describing roads for a given area (read in when the time is reached) and a buffer whose width is determined by the road gravity control factor, and defining the road attractiveness for development.
- *Seed layer* The initial distribution of urban areas that act as growth centers. The seed layer can be any distribution taken either from an actual time period or from a hypothetical starting distribution.



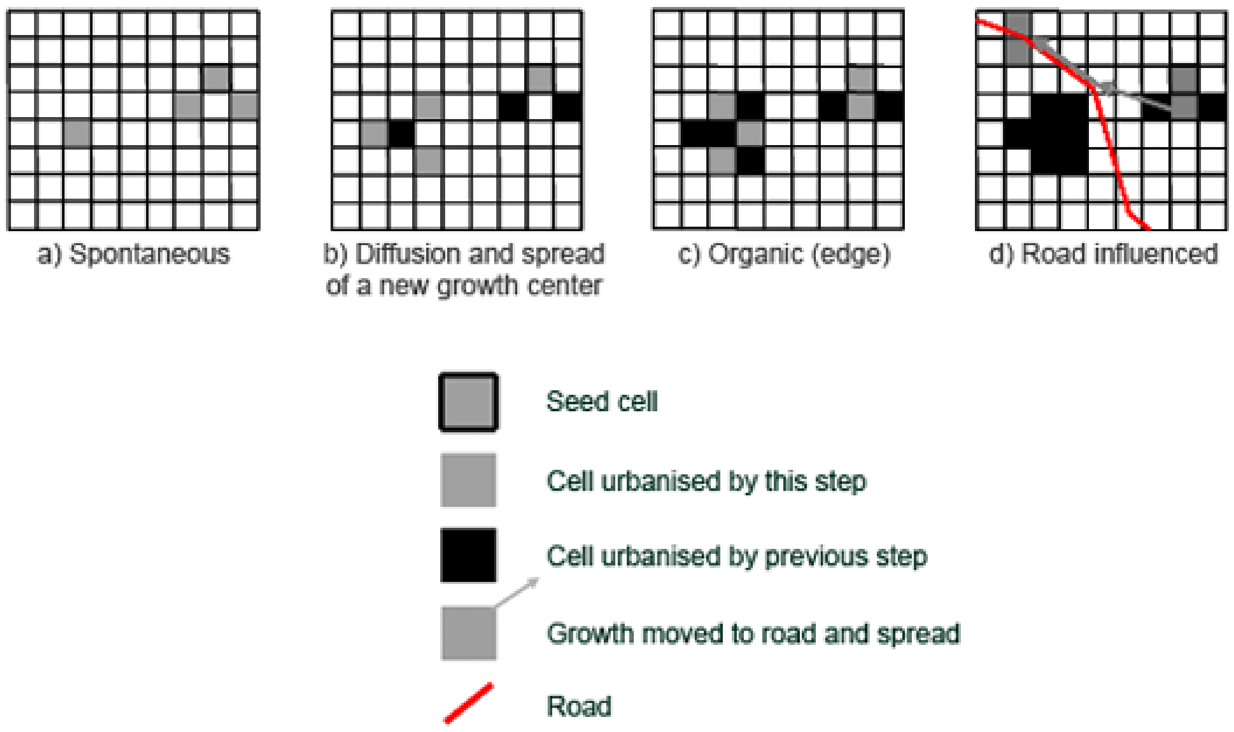
The four input data layers used by SLEUTH model for simulating urban growth (after Clarke et al., 1997)

[18] The SLEUTH name is the acronym for Slope, Landuse, Exclusion, Urban, Transportation and Hillshading. The Sleuth model is a cellular automata simulation model developed to model the development of urban areas (Clarke et al., 1998)

There are five factors that control the behaviour of the system, which are growth control parameters in the model:

- DIFFUSION, determines the overall dispersiveness of the distribution both of single grid cells and in the movement of new settlements outward from the road network
- BREED coefficient, determines how likely a newly generated detached settlement is to start its own growth cycle
- SPREAD coefficient, controls how much outward organic expansion takes place within the system
- SLOPE-RESISTANCE, influences the probability of settlements extending upward on steep slopes
- ROAD-GRAVITY, affects the attractiveness of new settlements onto the existing road network if they fall within a given distance of a road

The values for DIFFUSION, BREED, SPREAD, and SLOPE-RESISTANCE range from 0-100, and ROAD-GRAVITY ranges from 0-20.



Growth rules: model operation for a single cycle, a year for example (after Clarke et al., 1997)

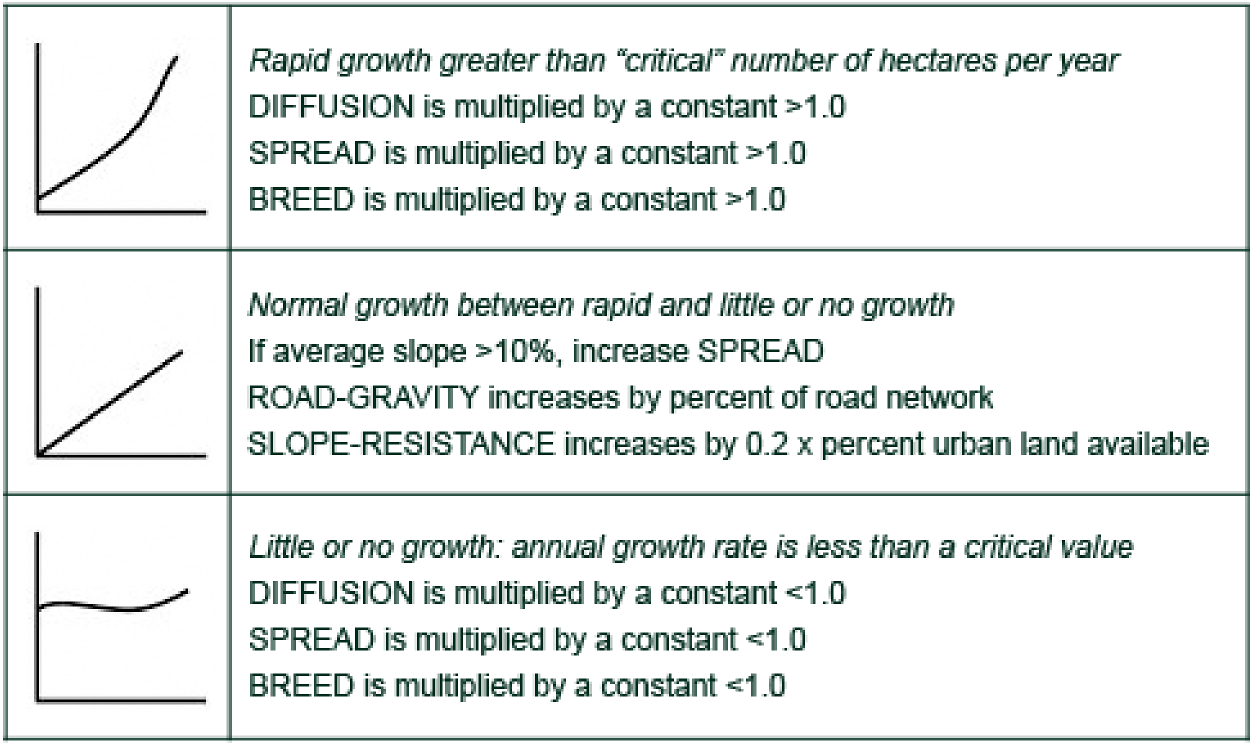
As we can see from the last figure illustrating the four growth rules, the model defines the growth rate as the sum of the four types of urban growth, which are:

- *Spontaneous growth:* it occurs when a randomly chosen cell falls in an area close to an already urbanised cell
- *Diffusive growth:* it urbanises cells that are flat and suitable for development even though they are not close enough to an urban area
- *Organic or edge growth:* spreads outward existing urban centers
- *Road-influenced growth:* promotes urbanised cells to develop along the road network

To allow the model to modify itself, a new set of rules are defined by coupling the variables. These modification rules can be summarised in four conditions:

1. When the absolute amount of growth in any year (cycle) exceeds a critical value, the DIFFUSION, SPREAD, and BREED factors are increased by a multiplier greater than one
2. When the system growth rate falls below another critical value, the DIFFUSION, SPREAD, and BREED factors are decreased by a multiplier less than one
3. The ROAD-GRAVITY factor is increased as the road network expands
4. The SLOPE-RESISTANCE factor is increased as the land available for development decreases

The following figure illustrates the self-modification rules.

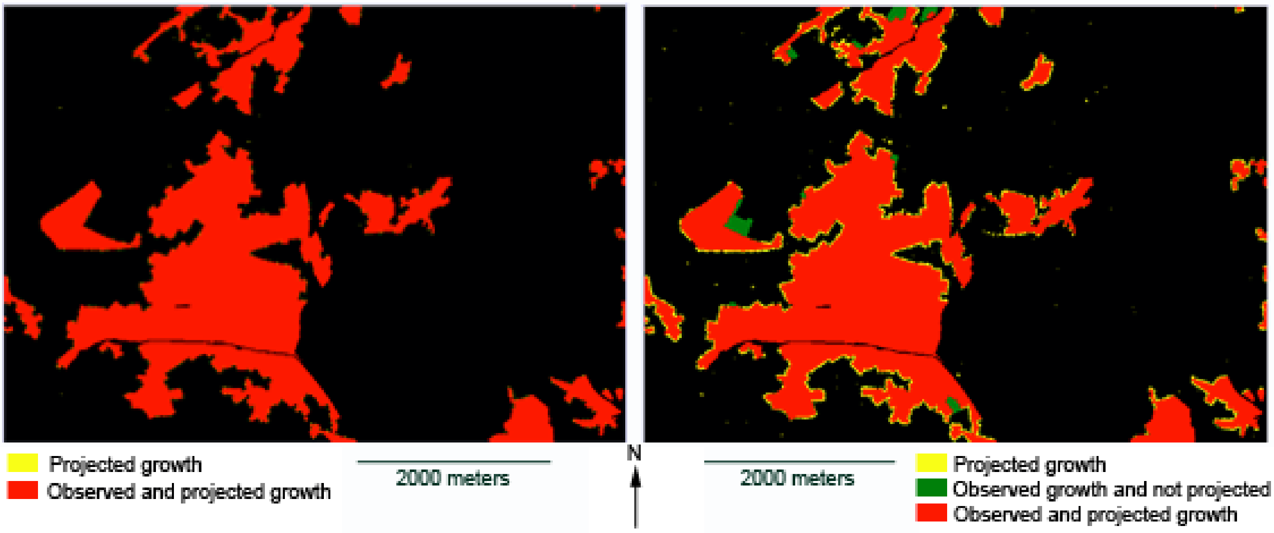


Self-modification adjustments to the control parameters (after Clarke et al., 1997

The SLEUTH model was applied to the study area of Bulle (Al-Ghamdi 2008) in order to simulate the urban development until the year 2050 according to three *scenarios of development* [19] with different levels of controlled expansion: extreme, moderate and less control scenario.
This development simulation requires two main stages: the calibration stage used to train and to evaluate the model capability based on past and documented LUC for the years 1952, 1974, 1993 and 2001, followed by the prediction stage that evaluates the urban extent for future and undocumented dates.

[19] A scenario of development describes the possible evolution of properties of a phenomenon, based on specific rules of change and specific characteristics (parameters) at the starting date. A set of different scenarios can be defined in order to simulate possible alternatives of development

Calibration stage [20]: five following layers that express the input factors controlling the urban growth process were introduced for the control years 1952, 1974, 1993 and 2001: LUC, urban extent, road network, excluded areas and slope. During this stage a validation phase was realised in order to estimate the capabilities to model the urban extent in 1993 and 2001. This model validation indicated a very satisfying simulation ability with a respective Kappa value of 0.98 and 0.95. The spatial distribution of projected and observed urban extent for 1993 and 2001 are illustrated in the next figure.



Projected and observed urban extent in Bulle area for 1993 and 2001 (after Al-Ghamdi, 2008)

Prediction stage [21]: mentioned above, this simulation is aimed to predict the urban development during the period 2001 to 2050. As the SLEUTH model authorises to modify rules and controlling factors during the iterative prediction stage, the road development plan for 2012 is added as an input layer as is the new planned protected areas layer (next table).

Layers	Control years				
Urban extent	1952	1974	1993	2001	
Land use			1993	2001	
Road network	1952	1974	1993	2001	2012
Protected area	Based on 2001 situation + new planned protected zones				
Slope	Constant				

Input layers used for the prediction stage (adapted from Al-Ghamdi, 2008)

[20] The application of a predictive model is made of 2 stages: calibration and prediction stages. Before a possible use for predicting properties in an undocumented period of time, it is necessary to calibrate (adjust, assign appropriate values to) the parameters of the model to suit the period of time validated with measurements

[21] The application of a predictive model is made of 2 stages: calibration and prediction stages. Once the model has been calibrated to suit the period of time validated with measurements, it can be used to predict feature properties estimated during a period of time with no existing measurements (generally a future period of time)

Three scenarios are submitted to forecast the urban development with different levels of controlled expansion: extreme, moderate and less control scenario. In the three scenarios all land use types are excluded from urbanisation, as having a control level value of 100 (**Table**), except for agriculture land that is left open if it is outside existing or new planned protected areas. A zero level of exclusion is then assigned to this land use type. The difference between the three scenarios is the following:

- *Extreme control scenario:* all existing and new planned protected areas are given a value of 100 which excludes them from urbanisation.
- *Moderate control scenario:* the new planned protected areas are assigned a value of 40 for the control level, giving more chance for urbanisation with moderate limitations.
- *Less control scenario:* the new planned protected areas are assigned a value of 0, making them open for urbanisation, depending on rules of growth.

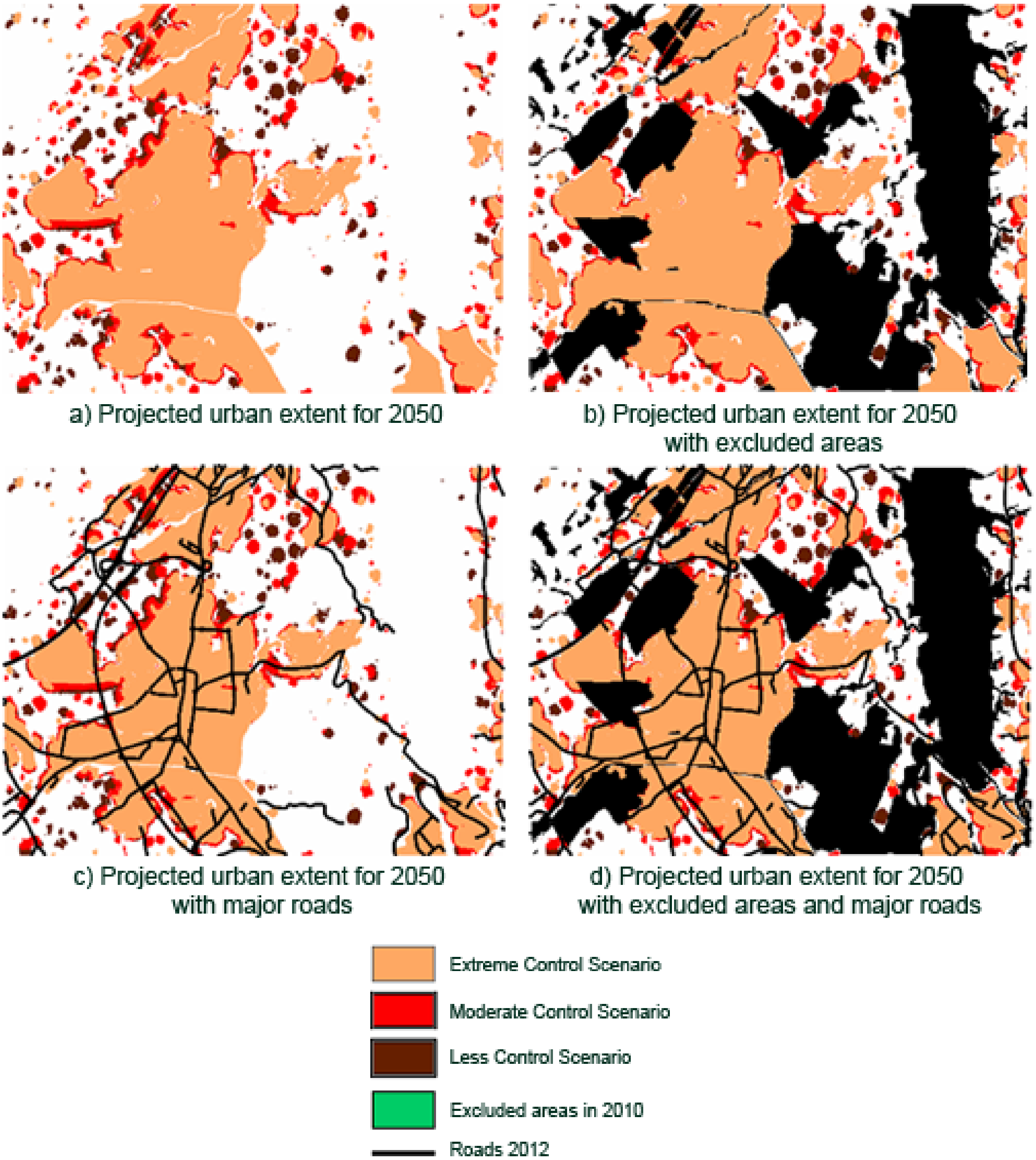
Scenario	Control level of exclusion (0 to 100%)					
	Water	Wetland	Forest	Parks	Agriculture	New protected zones
Extreme control	100	100	100	100	0	100
Moderate control	100	100	100	100	0	40
Less control	100	100	100	100	0	0

Three different scenarios and level of exclusion (adapted from Al-Ghamdi, 2008)

The next slideshow illustrates the urban growth simulation during the period 2001-2050 according to the conditions and rules assigned to the three scenarios.

Only pictures can be viewed in this version! For Flash, animations, movies etc. see online version. Only screenshots of animations will be displayed. [link]

Differences between the three scenarios at the end of the simulated period are illustrated in the next figure. One can observe the influence of the control level of exclusion on the spatial growth as well as the one from the road network.



Projected urban extent for 2050 according to the three scenarios and related with the road network and excluded areas (adapted from Al-Ghamdi, 2008).

EXERCISE:

From the last try to visually estimate:

- The impact of the control level of exclusion on the urban extent.

- The location of new urban areas and its relation with the road network.

1.4. Spatial Dynamics - Continuous case

When analysing the process of spatial growth for continuous phenomena, one should consider not only the *spatial expansion* [22] of a phenomenon in terms of presence or absence of its properties, but also the *change of intensity* throughout time. In each location of space, the intensity value can either increase or decrease at successive moments of the diffusion process. Therefore, for continuous phenomena, the binary concept of presence/absence can be replaced by another one called *densification* [23].

The concept of a phenomenon spreading through geographic space is considered in many diverse subject areas like the spread of wildfires, urban growth, infection diseases spread, diffusion of innovation, and ripple effects in the natural world.

In 1969, Peter Gould published a synthetic paper that clarifies the understanding of *spatial diffusion* [24] processes, with clear distinctions between diffusion types and definitions of basic concepts (Gould 1969) (Abler R. et al. 1972). Spatial expansion and densification can be modelled according to the following concepts. Spatial diffusion can be classified into three categories that represent the characteristics of the spread; we divide them in two groups: spatially dependant and non-spatially dependant diffusion.

1.4.1. Spatially dependent diffusion

In the *spatially dependent diffusion* [25] process, it is assumed that the spread is spatially continuous from one or several sources. *Contagious expansion diffusion* [26]: where the spreading phenomenon has a source and diffuses outwards into new contiguous areas, like wildfire or diffusion of innovation or an infectious disease.

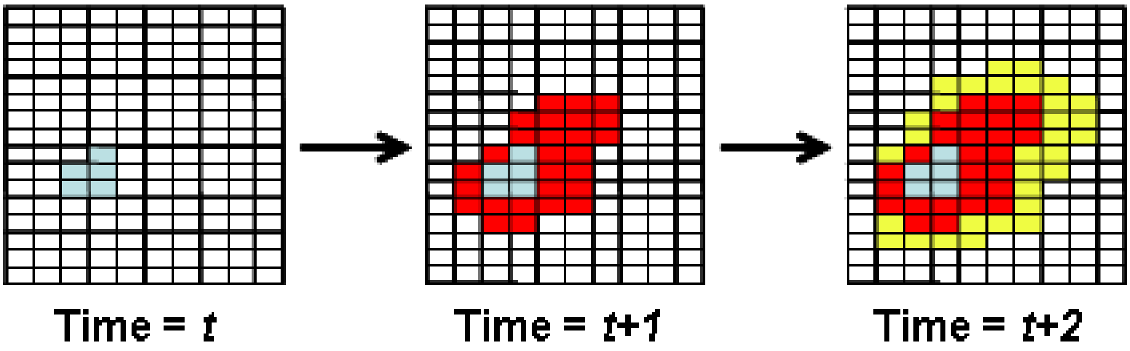
[22] The impact of the presence of a phenomeon throughout space. In a spatial diffusion process, the phenomenon spatially expands throughout time

[23] The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The densification expresses an increase in frequency of concerned items or individuals within a specific location or spatial unit of observation. In the process of diffusion of innovation, newly settled locations show an increase in the number of people adopting this innovation, this is known as densification (see Spatial diffusion)

[24] The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion)

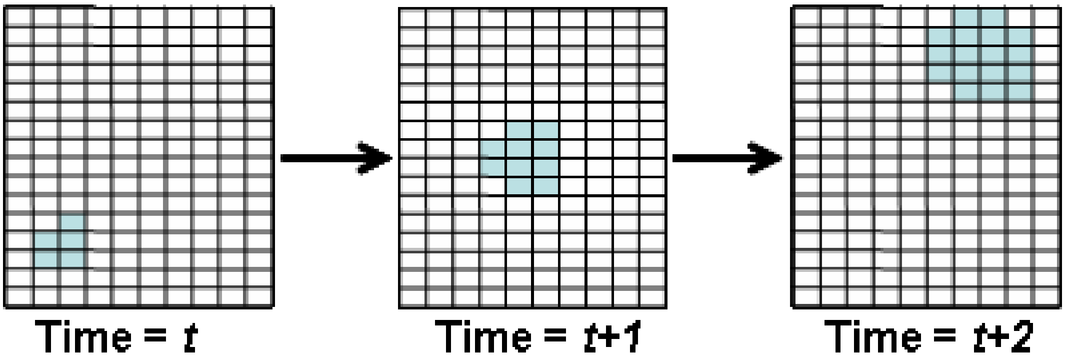
[25] The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion). In this second situation the diffusion process occurs with spatial continuity

[26] The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion). A contagious expansion diffusion is spatially dependant as the phenomenon spreads continuously from a location to contiguous neighbours, in a manner of a contagious disease



Principle of contagious expansion diffusion at successive moments of time

Relocation diffusion [27]: where the spread occurs when the spreading phenomena moves into new areas like migration. It could be interpreted as a movement or a travel in space. This specific spatial process will be considered in more details in the next **section**.



Principle of contagious relocation diffusion at successive moments of time. The feature is moving throughout space

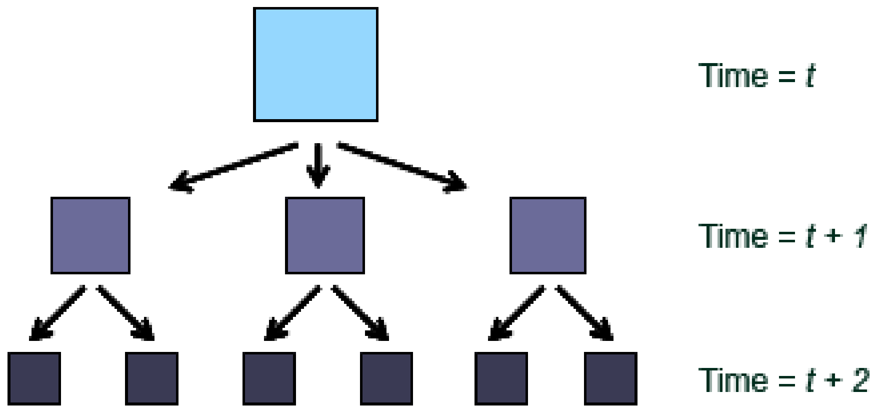
1.4.2. Non-spatially dependent diffusion

In the *non-spatially dependent diffusion* [28] process the spatial proximity is not influencing the behaviour of the diffusion. The major factor is the hierarchy of each place in space, such as the population size, or the economic, cultural or political influences. *Hierarchical diffusion* [29]: where the spatial diffusion occurs through an ordered sequence of classes or places like the spread of AIDS from large urban centres to small towns.

[27] A type of spatial diffusion process where previous locations of presence are replaced by new locations throughout time

[28] The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion). In this second situation the diffusion process occurs without spatial continuity

[29] The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion). In this second situation the diffusion process occurs without spatial continuity, but it is based on the hierarchical importance of centers (places) related with the innovative activity under investigation



Principle of non-spatially dependant Hierarchical diffusion at successive moments of time

Let us now concentrate on the spatially dependant diffusion process. We have seen that space can be modelled in different manners. If we admit that space cannot be considered as a homogeneous medium in the diffusion process, one should introduce a concept that characterises the specific influence of locations in the diffusion process. The concept of friction encompasses the overall specific properties of each location that influence the speed and the intensity of the diffusion process. In this way the spread can go through any place in space with more or less difficulties. The friction concept is expressing this level of difficulty.
In the theory of the spatial diffusion process, friction is considered as a barrier to the diffusion of innovation. At each location and for each moment during the diffusion process, the permeability level of a barrier can vary. Abler, Adams and Gould (1972) identify three different effects resulting from *barriers* [30] to the diffusion of innovation:

- Absorbing barrier completely blocks a pulse of innovation.
- Reflecting barrier will redirect the energy of diffusion toward different directions, such as a water body, for the expansion of a city.
- Permeable barrier absorbs part of the energy but allows the rest to go through. Its effects will slow down the process in its local area of influence.

Local factors that act as barriers to the diffusion process are of three types:

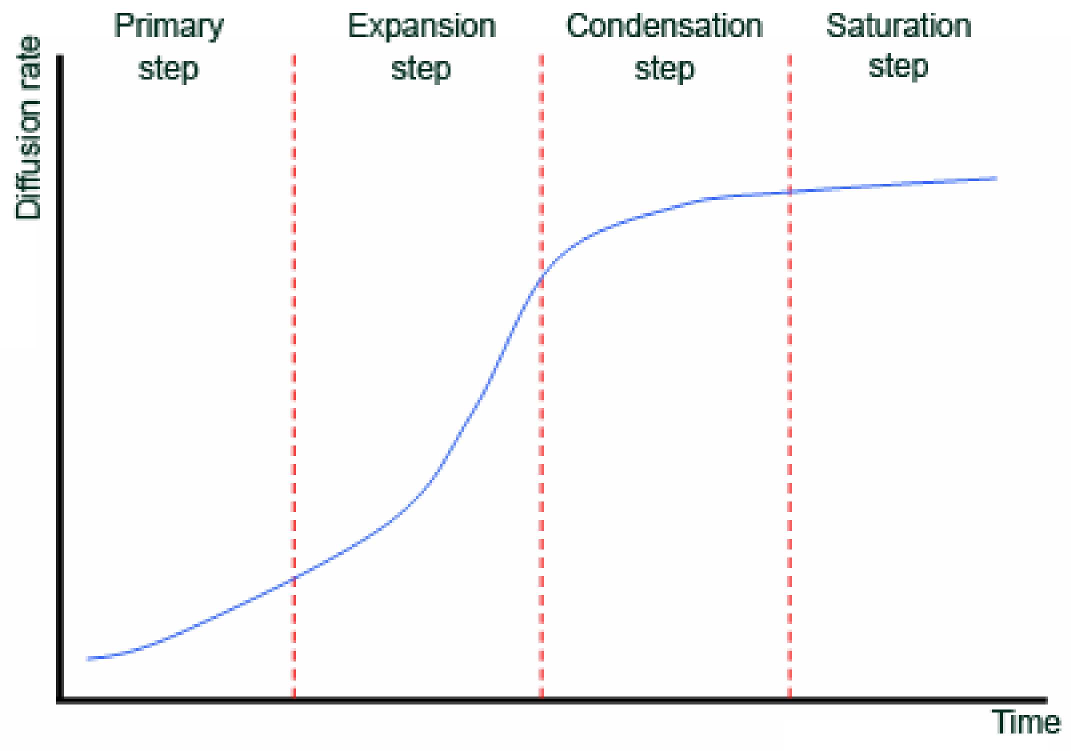
- Physical barriers that block or slow down the diffusion. They are physical properties of space such as the topography or the land cover.
- Cultural barriers can influence the diffusion of an innovation that spreads from individual acceptance. Linguistic, religious and political factors are typical cultural barriers to diffusion.
- Psychological barriers can be important for innovations involving individual acceptance in the process of diffusion. In this situation, individuals act as carriers in the diffusion process.

These concepts developed in the frame of the diffusion of innovation can be applied to the diffusion of numerous phenomena.
The temporal framing of spatial diffusion follows a number of rules:

[30] A barrier is a location that blocks movements due to its thematic property. It is often called "absolute barrier" and is modelled with an infinite friction coefficient value (see Friction)

1. Primary step of the process corresponds to the beginning of diffusion. At this stage, diffusion introduces a new differentiation inside geographical space; a contrast appears between places that have adopted the innovation and other places.
2. Expansion step is the stage of actual development of the process that generates a gradual softening of the strongest contrasts between places.
3. Condensation step, the rate of penetration into the different places tends to become more homogeneous, while speeds of diffusion in the various places grow closer.
4. Saturation step is the final stage where the diffusion rate increases toward a maximum following an asymptotic curve.

The diffusion rate change throughout time is generally modelled with a logistic function (S-shaped function) as illustrated in next figure.



Typical successive stages of diffusion rate during the diffusion process. For an innovation, they are called Innovators, Early majority, Late majority and Laggards

We now know most of the concepts and elements to understand the basic principles ruling spatial diffusion models, particularly those developed for the diffusion of innovation:

- The spatial expansion process is modelled with the component of friction that summarises the action of factors involved in the spatial interactions: distance, mass, barriers… Often contagious types of rules are combined with the hierarchical type, as both types are acting simultaneously.
- The densification in each location throughout time is controlled by the parameters of the logistic function.

Since the mid-sixties, when T. Hägerstrand had set-up the basis of the spatial diffusion model (Hägerstrand 1967), many authors and scientists have modelled the spatial diffusion of various phenomena such as technical innovations, new ideas, new cultural or social behaviours, but also the migration of persons or goods.

In his short paper, L. Frank (2002) clearly presents the application of a spatial diffusion model to the diffusion of mobile communications within the European Union. It is an interesting and illustrative application of spatial diffusion modelling.

1.4.3. Accessibility and movements in space

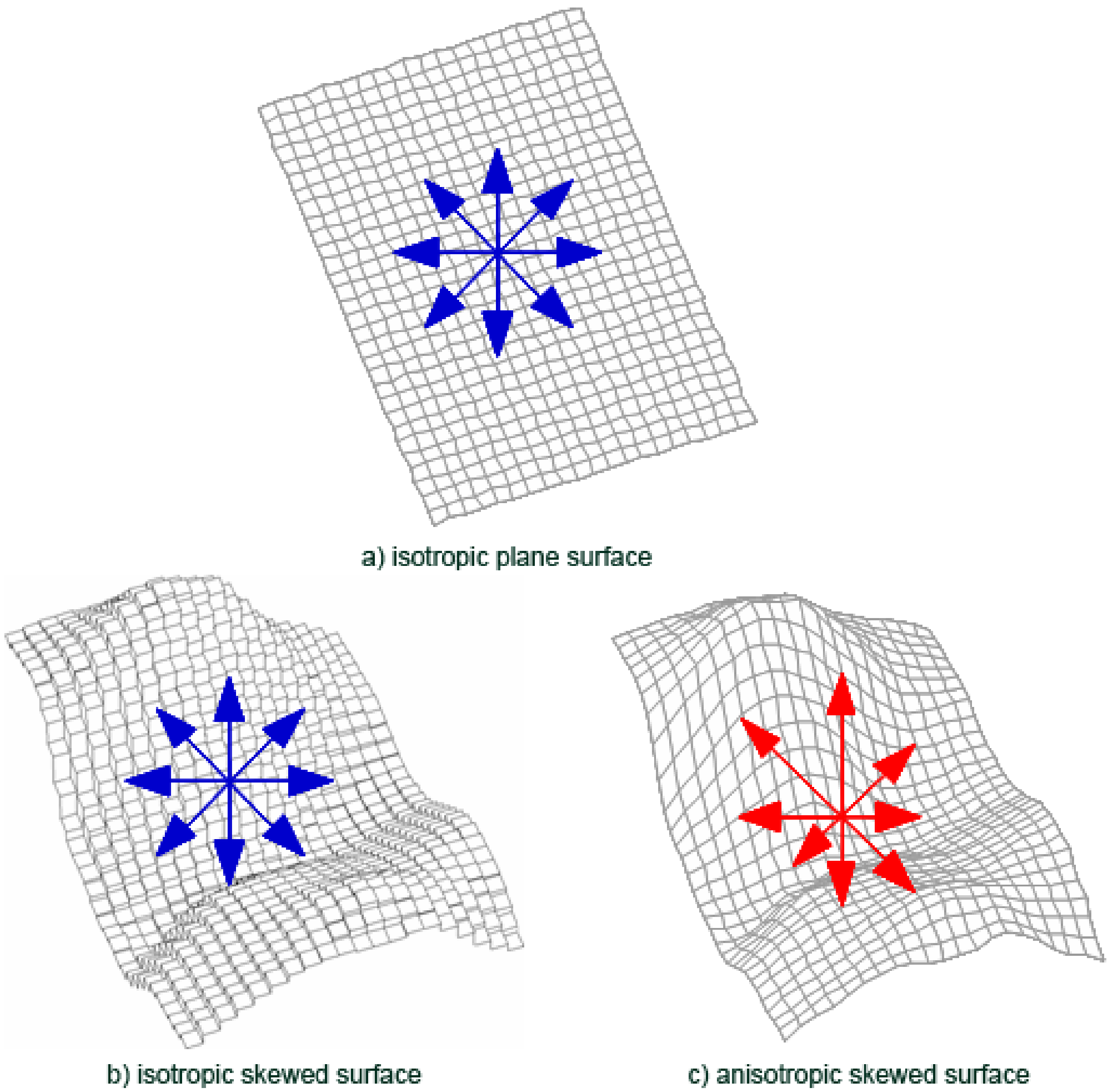
When combining the spatial and temporal dimensions, it is then possible to model trajectories in space. The modelling of trajectories concerns the type of moving features –people or goods- throughout space, with the means of transportation, as well as with the way space is modelled.
We already have identified three major levels of a model of space (**I-AN, 1. Introduction to Intermediate Spatial Analysis**, Section 1.2), from a "simple" homogeneous surface to a heterogeneous surface with anisotropic properties:

- As an *isotropic plane surface* [31]: space is simply considered as a homogeneous surface with thematic property distribution only ruled by Euclidian geometry (linear plane distance influencing accessibility, proximity and dependency).
- As an *isotropic skewed surface* [32]: space is considered as a heterogeneous surface with each location influencing differently the distribution of thematic properties as well as the proximity and the accessibility. Space is modelled as a skewed surface expressing an individual "isotropic friction rate" at each location. Distance is therefore no longer linear but symmetrical.
- As an *anisotropic skewed surface* [33]: space is considered as a heterogeneous surface but with an individual "anisotropic friction rate" at each location. Distance is therefore no longer linear nor symmetrical.

[31] Space can be modelled with different degrees of complexity, such as an isotropic plane surface, an isotropic skewed surface, or an anisotropic skewed surface. The isotropic plane surface is the simplest way to model space as a homogeneous plane surface controlled by euclidian geometric properties

[32] Space can be modelled with different degrees of complexity, such as an isotropic plane surface, an isotropic skewed surface or an anisotropic skewed surface. The isotropic skewed surface is the second level of complexity to model space as a heterogeneous surface. As each location in space has a specific influence on movements, the euclidian distance concept is replaced with the cost-distance concept

[33] Space can be modelled with different degrees of complexity, such as an isotropic plane surface, or an isotropic skewed surface or an anisotropic skewed surface. The anisotropic skewed surface is the third level of complexity to model space as an heterogeneous surface. Each location in space has a specific influence on movements, but this influence is not constant, instead it varies with respect to the relative direction of movement



Modelling spatial properties in different manners

When space is modeled with both geometric and thematic dimensions, it is then possible to model movements and trajectories in this space in a more complex, but more realistic, manner. Spatial dynamics models share some common concepts. From already discussed diffusion models, we will retain two basic concepts for modeling *accessibility* [34] and *movements in space* [35]: *barriers* and *friction* [36]. In reality, space is experienced

[34] A concept describing the possibility and the difficulty to reach a specific location when moving throughout space

[35] Actions to move from one location to another through space. The moving feature can be activated by external or internal forces and its path is influenced by local frictions of space

as an environment with heterogeneous properties with respect to movement. Each place retains or favours a variable rate of movement with moving features. This is modelled with the concept of friction. In extreme situations movement is strictly restricted within a network, such as road networks. Any other location in this space is inaccessible, acting as an impermeable barrier to movement; its friction is infinite.
Spatial dynamics models share some common concepts. From already discussed diffusion models, we will retain two basic concepts for modeling accessibility and movements in space: *barriers* and *friction*.
In reality, space is experienced as an environment with heterogeneous properties with respect to movement. Each place retains or favours a variable rate of movement with moving features. This is modelled with the concept of *friction*. In extreme situations movement is strictly restricted within a network, such as road networks. Any other location in this space is inaccessible, acting as an *impermeable barrier* to movement; its friction is infinite.
Major concepts applied in movement models:

- *Forces of movement* [37]: Energies allowing the movement of mobile features. Forces can be internal to the feature, giving it autonomous mobility and great flexibility to choose the direction of movement. Movement of a feature can also be generated or influenced by external forces such as gravity, wind or more general attraction forces. They can be either directional (anisotropic) or non-directional (isotropic). Internal and external forces can be simultaneously present, but according to their relative influence situations they can be classified into two groups: active movement for features with important internal forces (auto-mobility) and passive movement for features mainly moved by external forces.
- *Frictions of space*: Specific properties at any location in space can slow down a movement at varying degrees of intensity. However, this concept of friction can be extended to any influence on the movement and can therefore incorporate the concept of force. Influence can be split into three situations: a slow down effect, no effect and an accelerative effect. The property of directional variability can then be added to these friction characteristics with variable positive or negative effects, depending on the model of space in use.
- *Temporal variability* [38]: Depending on the level of complexity of the selected model, the temporal variability of forces, friction, and also direction can be integrated into the modelling process. One should remember that we are dealing with dynamic processes in which effects are varying throughout time.
- *Cost distance* [39]: When properties of space are taken into account with the concept of friction, then the Euclidian distance that was used to express proximity within a plane surface, should be replaced by a broader concept called Cost distance. As opposed to the plane distance that uses metric units, the cost

[36] When modelling movement in space, impeding properties of space are expressed as frictions. Friction coefficient values can range from 1 (no friction) to a maximum (infinite value) corresponding to an absolute barrier. In the computation of cost-distance the assigned friction value is combined with the geometric plane distance. The concept of a friction coefficient can be extended in order to encompass the concept of force. In this case the friction value is less than 1, but greater than 0 (see force of movement)

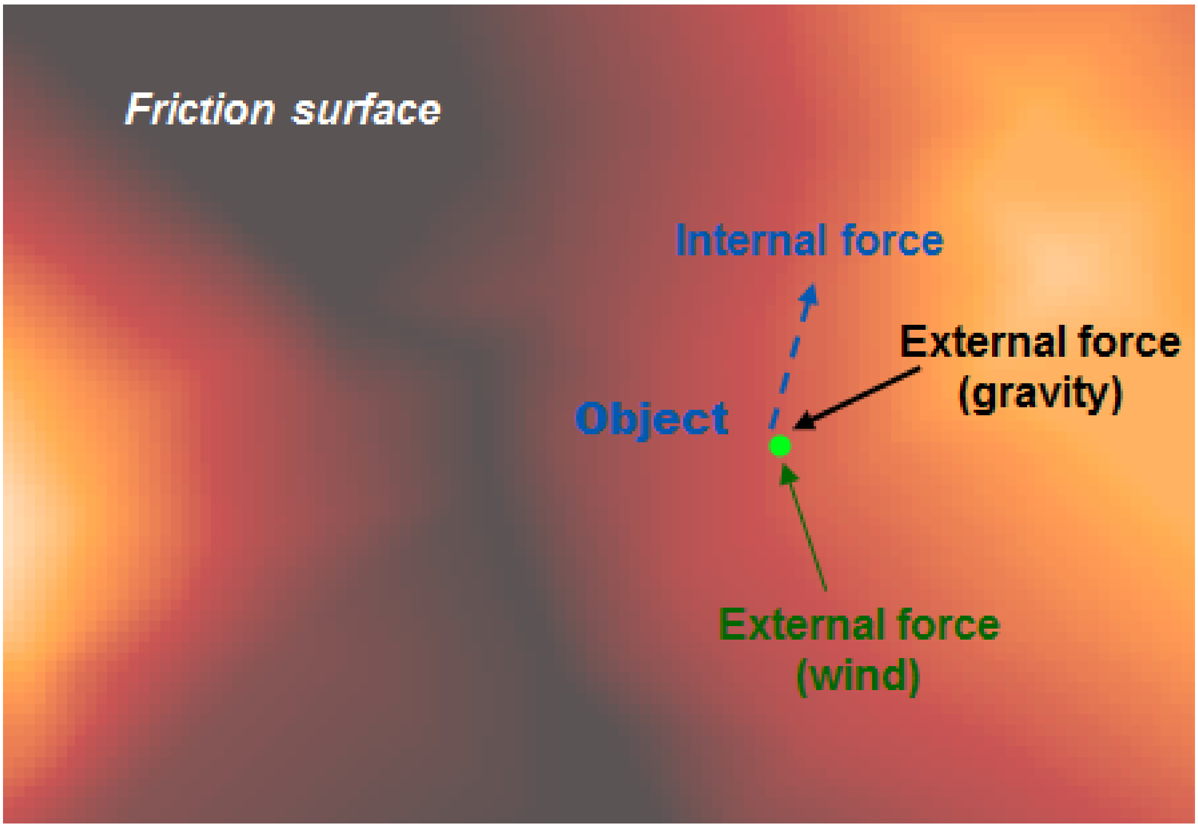
[37] When modelling movement in space, forces and frictions are factors that influence movement. Force can be internal to the moving body, but also external (wind, …). In the computation of cost-distance, the influencing forces are modelled as a friction coefficient value attached to each location in space. The friction value is less than 1, but greater than 0 (see friction)

[38] Some properties of space or spatial features can change throughout time, as others remain constant. Their temporal variability is then different

[39] When modelling access or movement in space, the usual euclidian distance -called also plane or horizontal distance- is not optimal to express the access to a destination. A most efficient concept is the cost-distance that express the amount of resources (financial, time, energy, ...) necessary to move throughout space (see Friction, Barriers)

distance expresses the amount of resources spent while moving from one place of space to another. Units of measurement can then express an amount of time, energy, financial resources, or any other type of resource. Cost distance is also called weighted distance, as the geometric distance is weighted by friction coefficients.

- *Path, trajectory* [40]: The modelled path or the trajectory of movement from one place to another corresponds to the itinerary that optimises the considered resource according to the friction surface and the active forces. It is often called the optimal path, but might not be unique, particularly in an open space modelled in a raster structure.



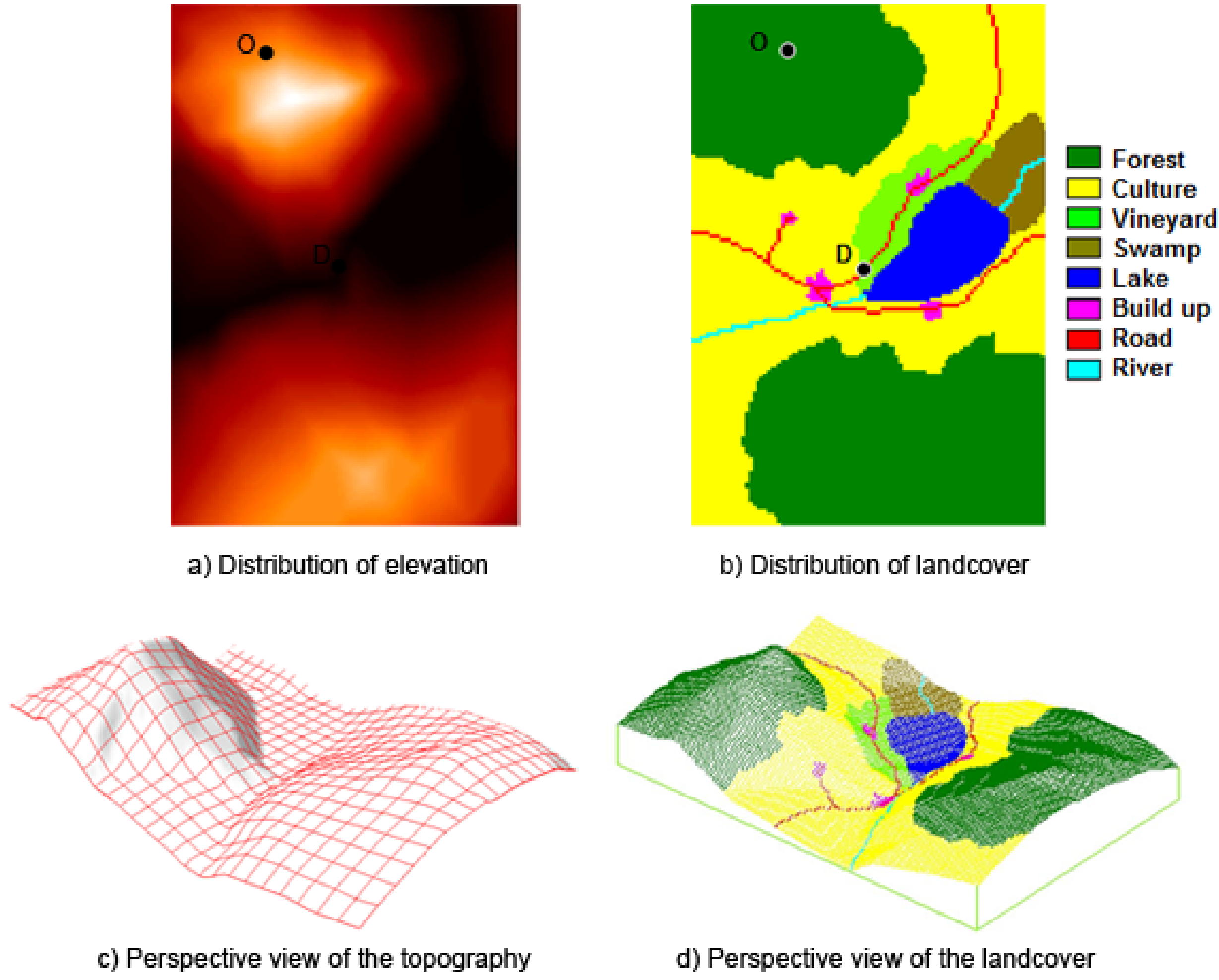
Major components influencing the movement in space

1.4.4. Illustration with the movement of a walker

For didactical purposes this section is organised as follows: Principles of accessibility modelling are presented in turn for the three levels of modelling space and its properties. A common example will illustrate for these three levels the movement of a walker. This specific situation is characterised by an active moving person with an internal force, by a moving strategy that is influenced by several factors related with individual properties of places in space.

This rather simple example should be familiar to most of us and is suitable for the three levels of models. Let us describe the situation of this illustration. A person wishes to walk from a starting place –an origin- to a destination within our imaginary region of Examplis. Its topography is moderate with a heterogeneous landcover. Next figure illustrates the topography and the landcover of this region and shows the origin (O) and the destination (D) of the walker's movement.

[40] The trajectory used to move from an origin to a destination in space



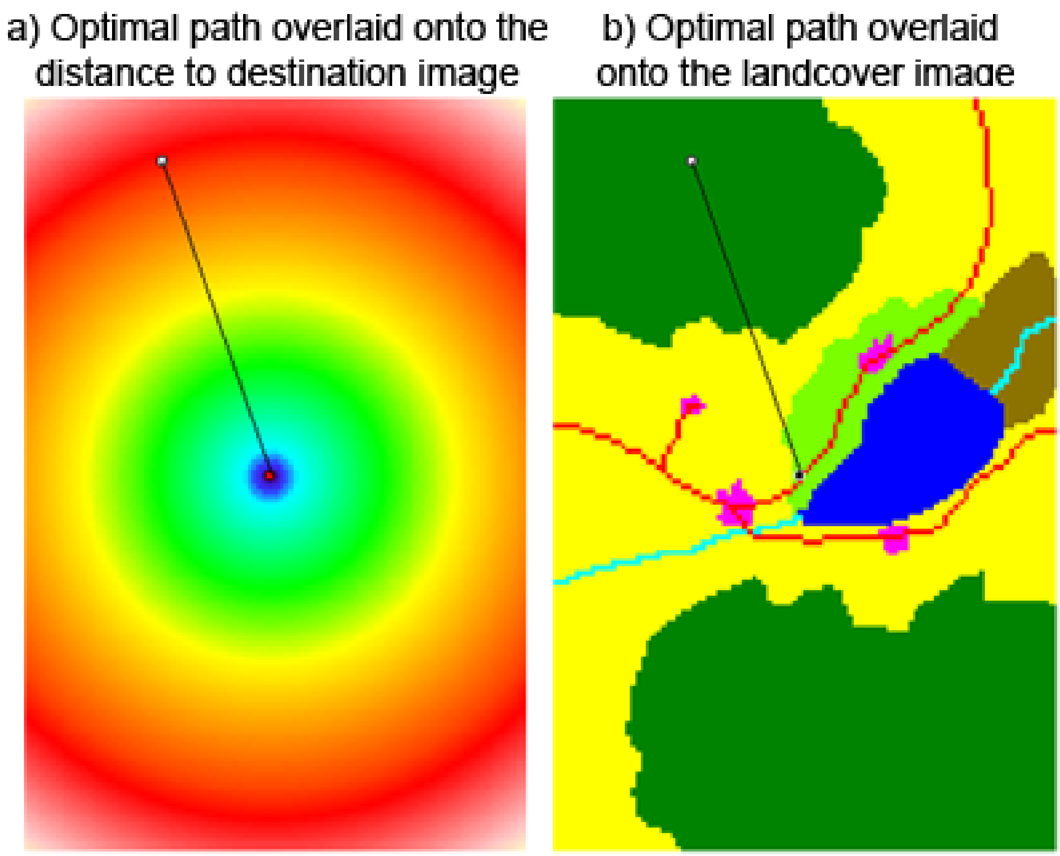
Topography and landcover distributions in the region of Examplis. The origin (O) and the destination (D) of the walker's movement are indicated

Modeling movement in an isotropic plane surface

In this modeling context, only geometric properties are influencing the movement of the walker. The measure of distance between locations in space, their proximity and the length of the path are expressed with metric units. Principal characteristics of the model at this level are the following:

- The distance to the destination increases linearly and homogeneously in all directions as a function of the proximity to this location. These two characteristics of regularity and isotropy correspond graphically to a concentric distribution of distances with maximal values on the corners of the region (figure 3.26a).
- According to the Euclidian geometry rules, the optimal path between the origin and destination is a line linking them.

We can observe on next figure a) that space is considered as homogeneous in this model and therefore only the geometric distance is considered for the estimation of proximity from any place to the destination. This is confirmed with the overlay of the optimal path onto the landcover distribution (figure b). One can observe that there is no influence.

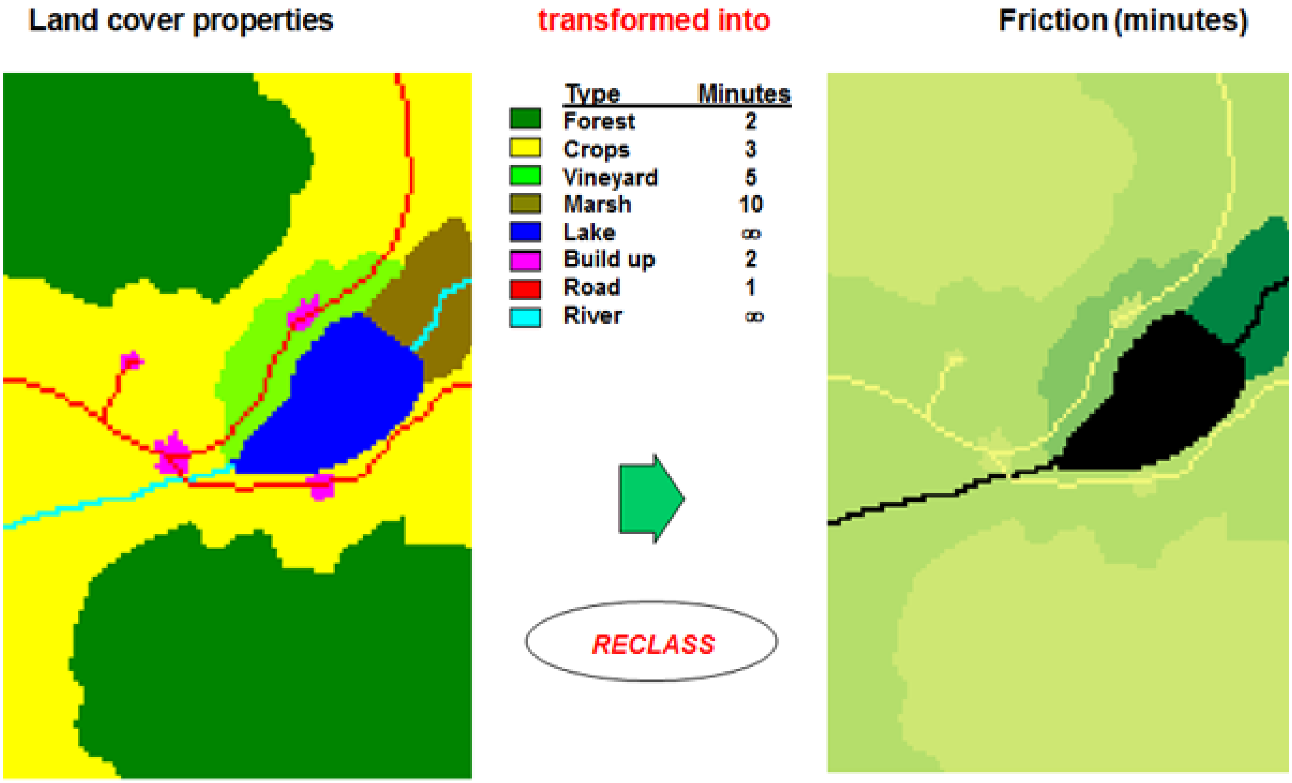


Distance to the destination and optimal path modeled within an isotropic plane surface

Modeling movement in an isotropic skewed surface

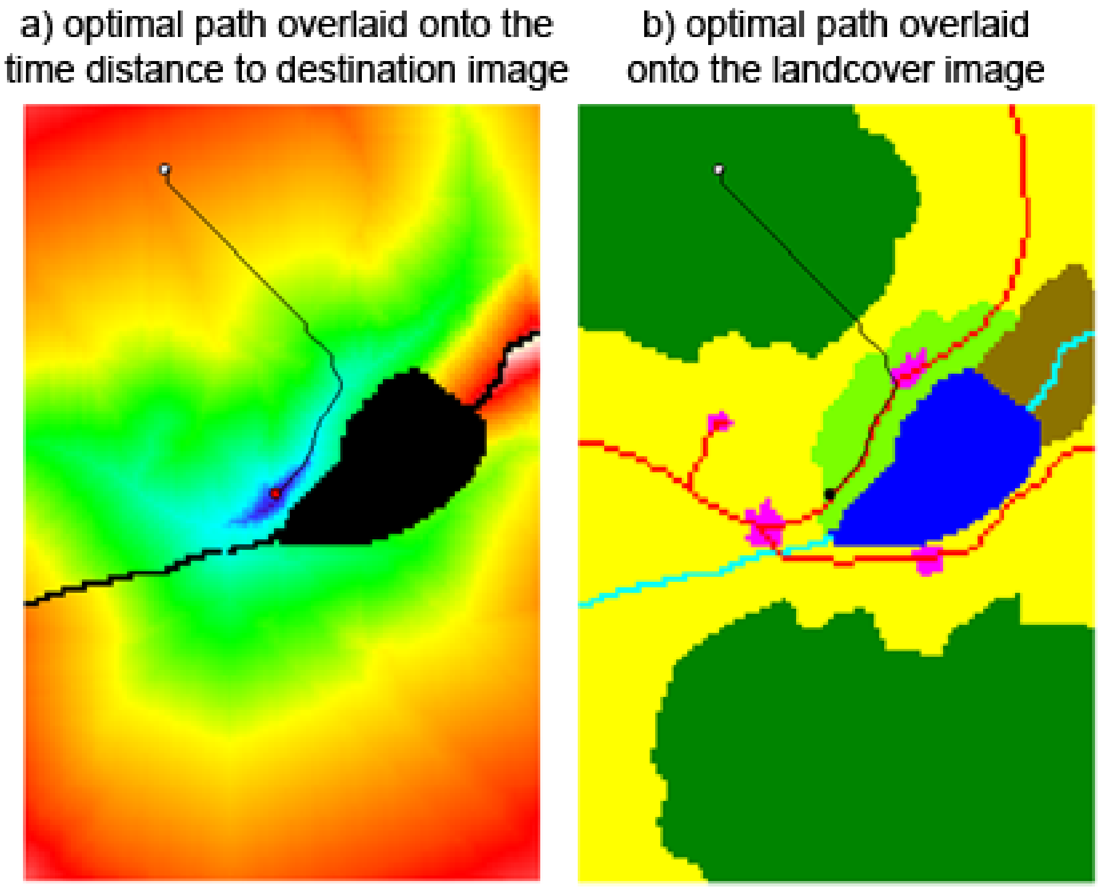
For this modeling context, it is possible to take into account both the Euclidian distance from the geometric dimension and the friction effect due to individual properties of space in the thematic dimension. It is then necessary to construct a layer of friction coefficients that express the friction effect of each cell on the movement of the walker.

In this application let us limit the consideration of a single friction factor to the movement: the cover at the surface of the land. Our concern is to translate the landcover properties in terms of friction effect. Let us choose a cost distance unit as the time unit in minutes and therefore call this distance a time distance. The coefficient friction value attached to each cell corresponds to the amount of time necessary to cross this cell according to its landcover friction effect. Assuming a cell size of 100m, the friction value assigned to each landcover type is listed in the next figure. The production of the friction layer illustrated in the next figure is based on a simple classification of landcover types into minutes. We can observe that the road cover has the lowest friction coefficient, the marshy cover a high value and water covers have an infinite friction value, corresponding to an impermeable barrier.



Production of a friction layer expressing the time to cross a cell. This layer is obtained by replacing (recoding) landcover type values with their corresponding assigned friction coefficient

The computation of the time distance from each cell to the destination cell can now be started through an iterative and cumulative process from the destination to all image cells. The GIS software Idrisi offers such a procedure called COST-GROW. The next figure a) shows the optimal path of the walker overlayed onto the time distance image. One can observe that the most distant location is no more than four image corners, like with the Euclidian distance, but the southern part of the river that flows through the marshy cover . In figure b) the optimal path has been overlaid onto the landcover image in order to illustrate the influence of the considered friction factor on the optimal path.



Optimal path for the walker modelled on the basis of the landcover properties as the friction factor

Modeling movement in an anisotropic skewed surface

In the third level of this model of space, the heterogeneity of friction properties can be taken into account. Let us consider the factor topography to illustrate the anisotropic friction effect. By experience we should admit that the slope of a cell has not the same influence when walking downhill or uphill. The slope steepness of that cell is identical, but its friction effect varies according to the direction of movement of the walker relatively to the direction of the slope. If we consider more carefully the effect of this topography factor we can observe that in direction of the steepest descending slope the friction effect is in fact transformed into a force –the gravity force- that aids to walker to move across the cell. In other words, this anisotropic characteristic allows us to consider friction factors for which friction varies as a function of the direction of movement. As you remember we have admitted that the influence of a factor acting either as a force or a friction can be modeled with a unique concept of relative friction. In order to model this influence we can consider a force-friction vector made of two components: its intensity and a function of directional variability.

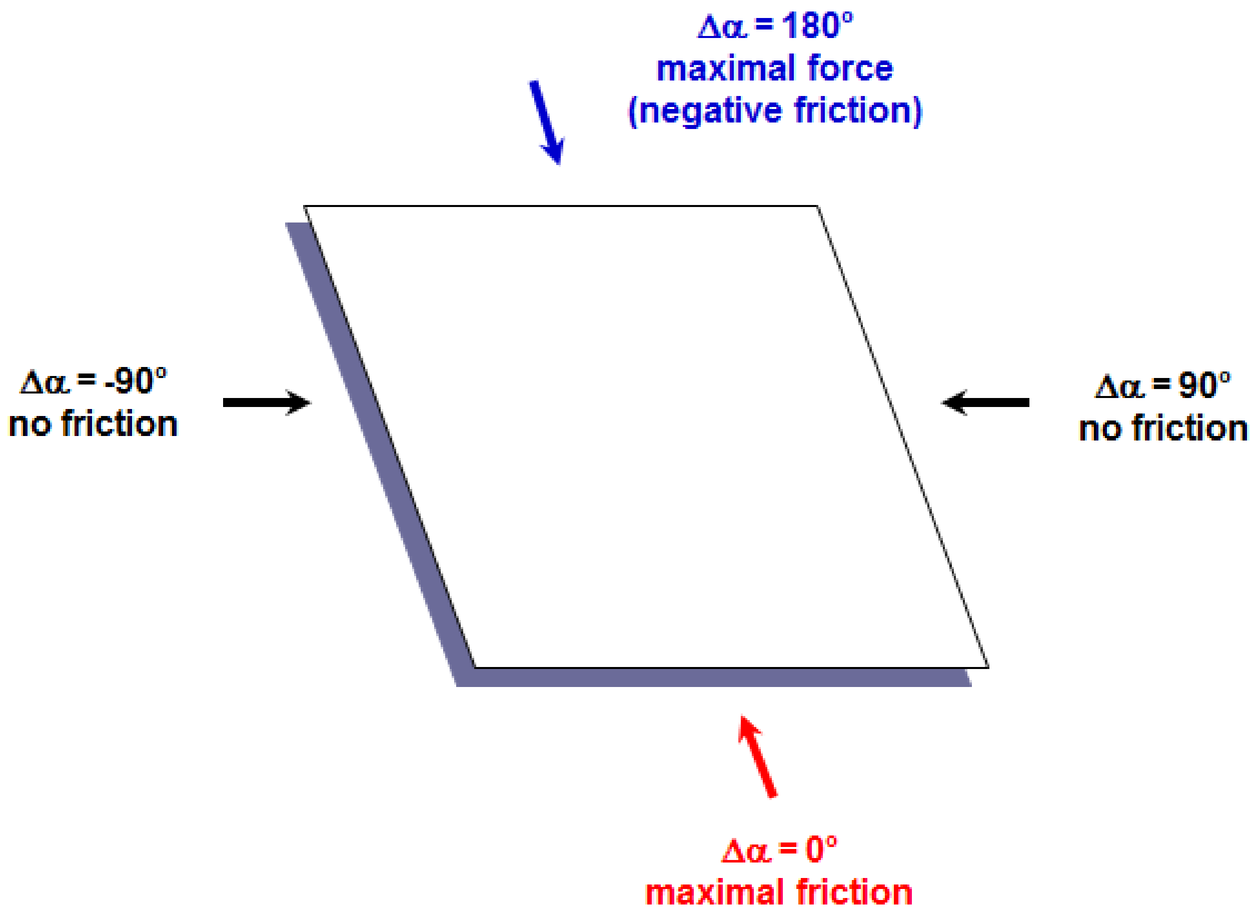
Let us return to our application of modeling the movement of our walker. Our objective is to consider the effect of both the landcover and the topography along with the walking time and pathway. Let us proceed in two steps, first we concentrate on the modeling of the topographic factor and in the second step we combine its joint influence with the landcover factor previously discussed.

The factor topography with anisotropic properties The force-friction effect on the walker's movement is a function of the direction of walking with respect to the direction of the slope on each piece of terrain: a cell. The next figure illustrates the strength of this effect for different relative directions noted ##. We can then identify three typical situations:

- Up hill situation: When the direction of movement matches the direction of the steepest slope, the relative angle ## = 0° and therefore the friction effect of the slope is maximal.
- No slope situation: When the direction of movement is perpendicular to the direction of the steepest slope, the relative angle ## = 90° or 270° and the slope effect is neutralized. Therefore the friction effect of the slope is unitary (=1), corresponding to a flat surface.
- Downhill situation: When the direction of movement is opposite to the direction of the steepest slope, the relative angle ## = 180° and the gravity force effect of the slope is maximal, corresponding to a minimal relative friction

We have then identified three key situations with three corresponding relative friction values:

- A maximal friction value corresponding to the required time distance for climbing this slope.
- A unitary friction value expressing a neutralized influence of slope. In this situation the time-distance is only influenced by the length of the cell to cross: its geometric distance.
- A minimal friction value corresponding to the influence of the gravity force on the time distance for descending this slope. This minimal friction value is a decimal value ranging between 0 and 1, but always greater than 0 and less than 1, depending on the effect of that force.



Anisotropic impact of the slope factor on the movement of the walker crossing cells. Depending on the relative direction of the movement with respect to the direction of steepest slope ##, the relative friction effect range from a maximal friction to a maximal force

How do we determine the relative friction value for intermediate situations? We can admit that the friction effect of slope should move continuously between the identified minimal and maximal values and should be equal to 1 for a flat slope. We can then model this anisotropic effect of friction with the use of a continuous function related with the relative direction of movement ##. We can observe from the figure 3.29a) that the friction effect is symmetrical with respect to ##. Therefore only anisotropic function coefficients in the range of 0° to 180° should be estimated. In fact this *anisotropic function* [41] results from a combination of two functions: one describing the decrease of the friction effect from the maximal value to the unitary value, in the range of ## = 0° to 90° and the second that expresses the increase of the force effect from the unitary value to the minimal value, in the range of ## = 90° to 180°. The modelling procedure proposed in Idrisi called VARCOST offers several options for constructing this anisotropic function that will lead to the estimation of a relative direction friction value. One should keep in mind that the definition of the anisotropic function determines the rate of change of the friction effect as a function of the relative direction. This of course depends on the nature of the factor to model and the choice is left to the decision of the analyst.
In order to estimate the relative friction (called effective friction in Idrisi) of a slope for any ## relative direction value, the maximal friction value can be combined in the following form:

[41] A function that models the change of influence of a feature property with respect to its relative positional direction. This function is used to weight the friction/force effect (see Friction, Anisotropic skewed surface)

$$f_{reff} = f_{rmax}{}^{f(\Delta\alpha)}$$

With:

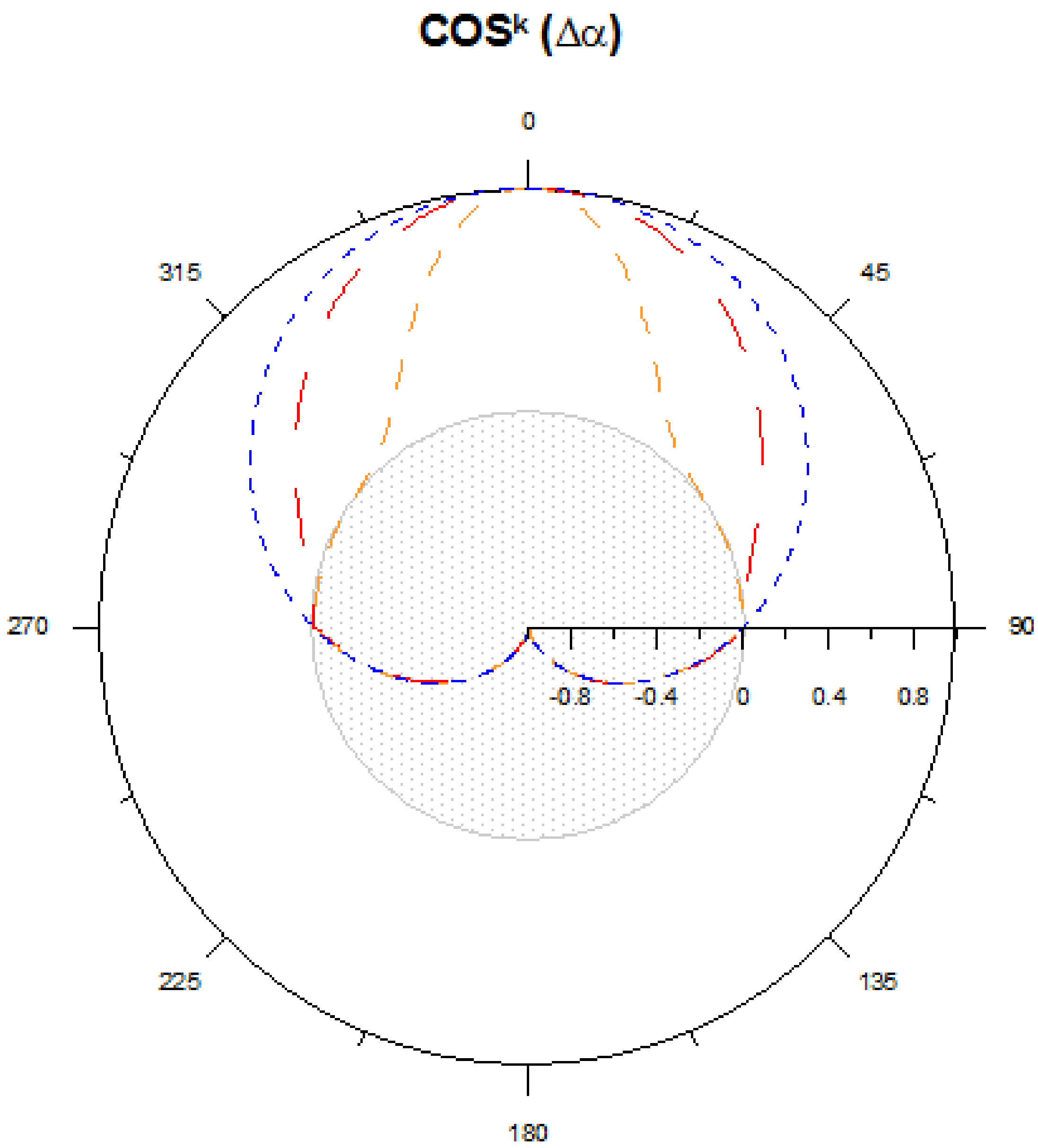
f_{reff} : effective friction coefficient value for a given relative direction of movement $\Delta\alpha$

f_{rmax} : maximal friction coefficient value assigned to the slope value

$f(\Delta\alpha)$: anisotropic function with values varying in the range 1 to -1

Let us illustrate the construction of an anisotropic function f() that gives a weight to the maximal friction to produce the final effective friction value. We decided to combine two different functions:

- A cosinus function cosk to model the decrease of the friction in the range 0° to 90° symmetrically from 0° to 270°. Values of the function should vary from 1 to 0 in order to produce an effective friction (freff) varying from its maximum value (frmax) to a friction value of 1 (ie. freff = frmax0 = 1). As illustrated in the next figure, an increase in the exponent k produces a faster decrease of f values with respect to the angle values ##.
- A linear function to model the increase of the force in the range 90° to 180° and symmetrically from 270° to 180°. Values of the anisotropic function should vary from 0 to -1 in order to produce an effective friction (freff) varying from 1 to a minimum value of 1/frmax (ie. freff = 1/frmax).

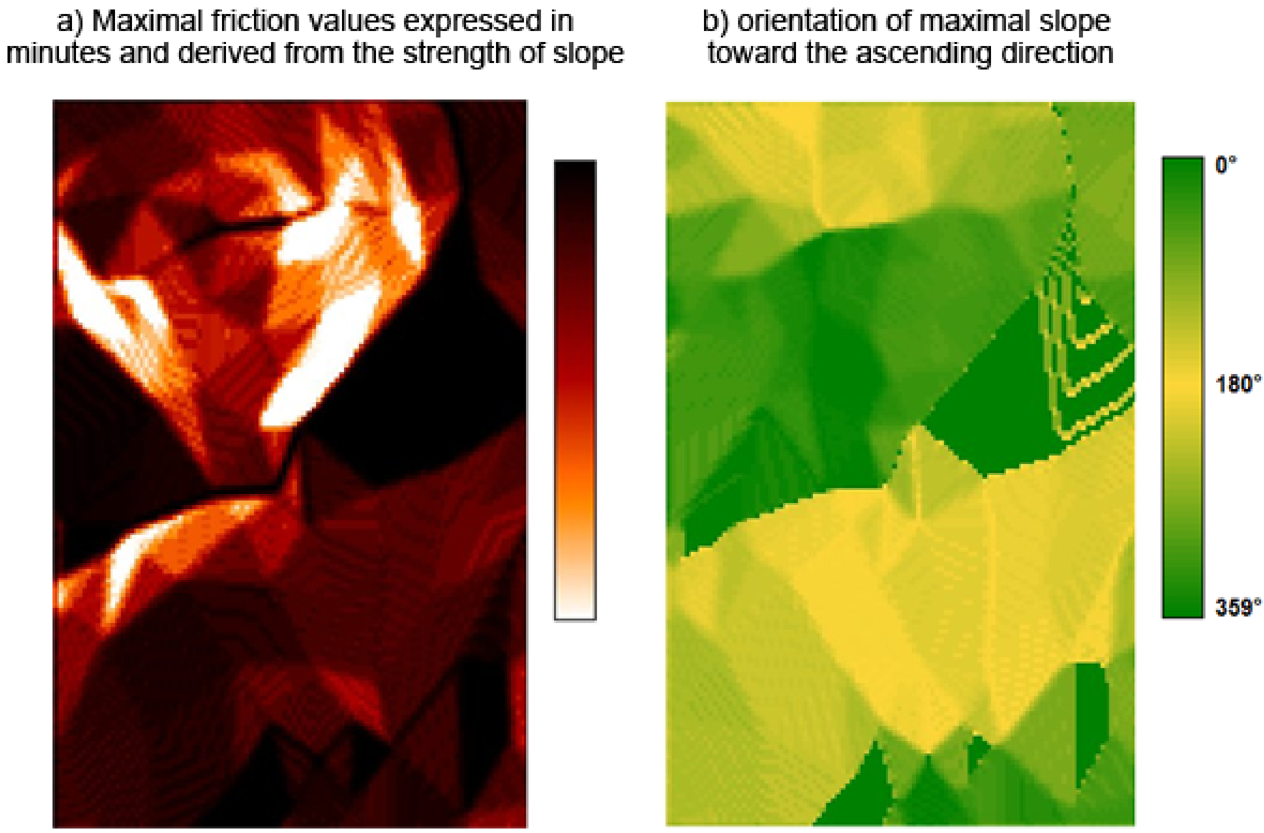


The anisotropic continuous function cosinus k controls the rate of change from a maximal friction to a maximal force. The k exponent acts on the strength of decrease from a maximal friction to a unitary friction (blue: k=1, red: k=2, orange: k=10)

The chosen simulation procedure VARCOST requires three input components:

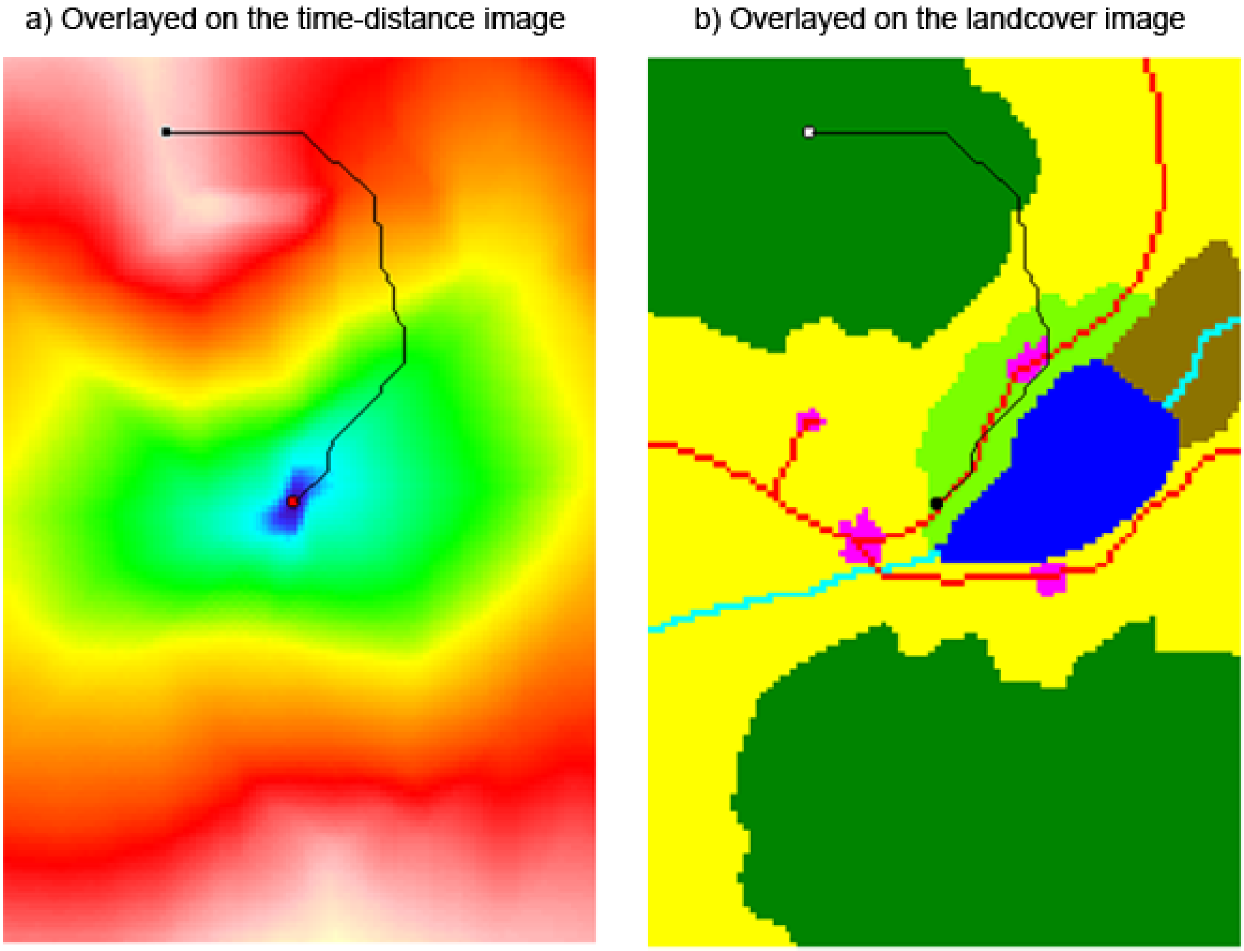
- An image layer of maximal friction coefficient that expresses the cost distance for crossing each cell based on its maximal slope.
- An image layer of the maximal slope direction (aspect) that expresses the upward orientation of the maximal slope. In practice, most procedures that derive the maximal slope orientation from a DEM will compute it as the descending direction. It is therefore necessary to transform it into an ascending orientation that corresponds to the direction of the maximal friction.
- An anisotropic function that weight the maximal friction index according to the relative direction of movement.

Let us first assign a friction coefficient to slope values. In our example we want to express a time distance of movement. Eastman (2008) is proposing a formula transforming the slope amplitude into time of movement. This transformation produces our maximal friction coefficient layer, as illustrated in the next figure a). The corresponding ascending orientation layer is presented in b).



The two required image layers for simulating the anisotropic effect of the slope factor on walking movement

Based on these two image layers and with the used of the anisotropic function cos1, the time-distance image to the destination D can be generated, along with the optimal path from the origin O, as illustrated in the next figure a). Figure b) shows that the optimal path does not match the landcover distribution, simply because it was not considered during this simulation stage.

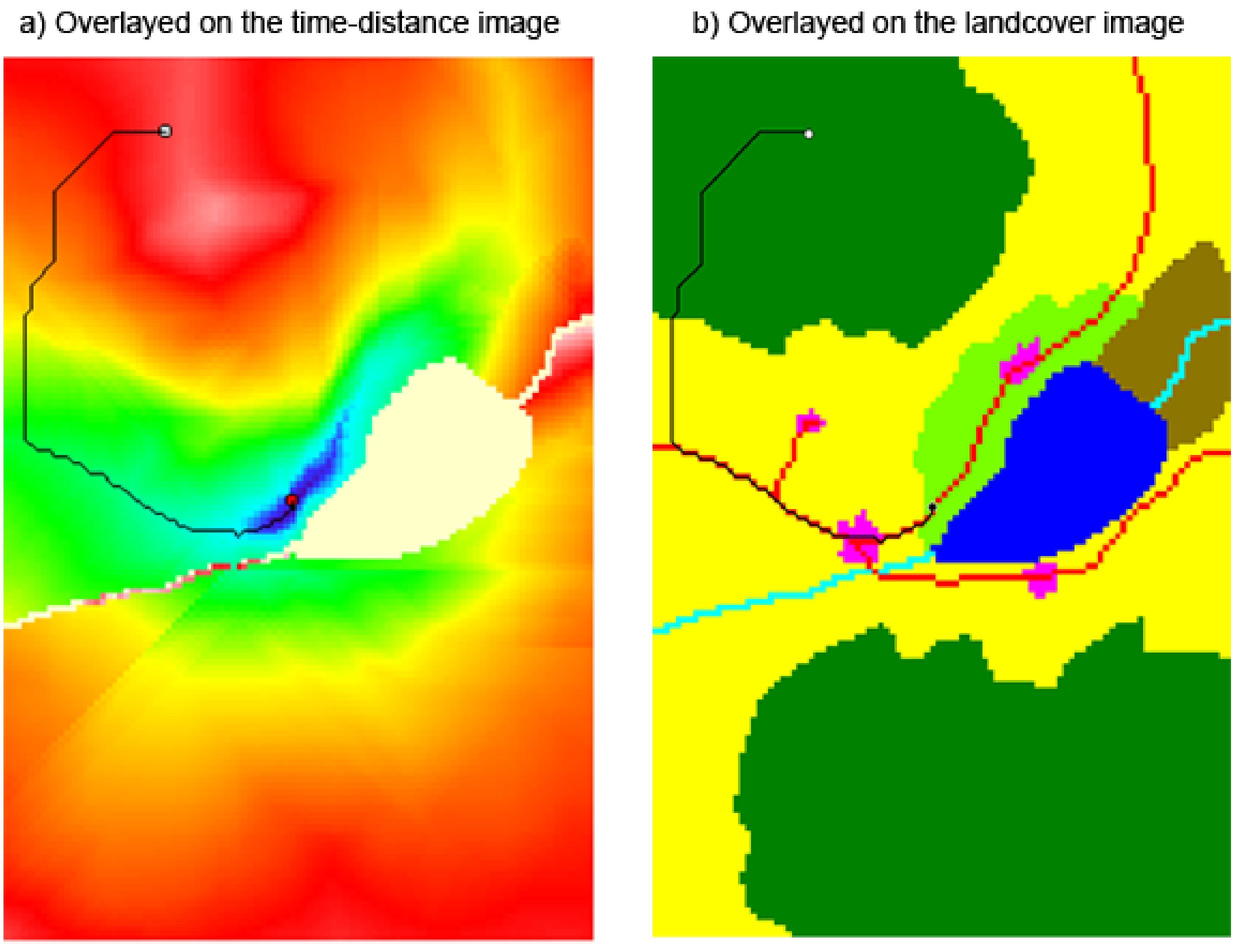


Optimal path of the walker modeled on the basis of the anisotropic influence of the topography

Combination of the two factors: topography and landcover

Let us now move to the ultimate stage of our simulation modeling by taking into account the simultaneous influences of the topography and the landcover on the walking movement. The procedure VARCOST allows us to combine several friction factors in order to generate a resulting friction layer and then the final time-distance image to the destination.

The next figure a) shows the resulting time-distance image with the optimal path overlayed, while figure b) illustrates the relationship of the pathway with the landcover. One can observe the added effect of the landcover friction on the optimal path: the hill is now walked from the west side and it reaches the nearest road location as fast as possible because of the low friction attached to this latter cover.



Optimal path of the walker modeled on the basis of the combined influence of the topography and the landcover

1.5. Summary

The combination of the time and spatial dimensions permits the performance of two distinct types of change analysis: the change in the patterns of distribution of the properties of spatial features and the simulation of movements in space.
This Unit presents several methodologies to investigate these two approaches. The spatial change analysis concentrates on how space evolves throughout time by the identification of local and regional change components. It makes use of the change indices presented in **lesson 8** and explores the spatial distribution of these changes in terms of intensity and direction of change. On the other hand, the modelling and simulation of movements in space deal with the process of space accessibility. Cellular automata and diffusion models offer tools for analysing spatial growth processes, whereas accessibility processes can be analysed with the use of concepts such as the cost distance, frictions, barriers, and anisotropic properties of space.

1.6. Recommended Reading

- **Abler R., Adams J., Gould P.**, 1972. *Spatial Organization, The Geographer's View of the World*. USA: Prentice Hall.
- **Caloz R., Collet C.**, 2011. *Analyse spatiale de l'information géographique*. Lausanne, Switzerland: PPUR.
- **Davis, J.C.**, 1986. *Statistics and data analysis in geology*. New York: John Wiley &Sons.
- **Eastman, R.**, 2008. *IDRISI - Taiga GIS and Image processing software, Reference Manual*. Worcester, Clark Labs, Clark University, USA: Worcester.

1.7. Glossary

Accessibility:

A concept describing the possibility and the difficulty to reach a specific location when moving throughout space

Anisotropic function:

A function that models the change of influence of a feature property with respect to its relative positional direction. This function is used to weight the friction/force effect (see Friction, Anisotropic skewed surface)

Anisotropic skewed surface:

Space can be modelled with different degrees of complexity, such as an isotropic plane surface, or an isotropic skewed surface or an anisotropic skewed surface. The **anisotropic skewed surface** is the third level of complexity to model space as an heterogeneous surface. Each location in space has a specific influence on movements, but this influence is not constant, instead it varies with respect to the relative direction of movement

Barrier:

A barrier is a location that blocks movements due to its thematic property. It is often called "absolute barrier" and is modelled with an infinite friction coefficient value (see Friction)

Calibration stage:

The application of a predictive model is made of 2 stages: **calibration** and prediction stages. Before a possible use for predicting properties in an undocumented period of time, it is necessary to calibrate (adjust, assign appropriate values to) the parameters of the model to suit the period of time validated with measurements

Cellular automata:

Cellular automata are dynamic systems used to model spatial changes such as spatial diffusion, movements, … Cellular automata are dynamic agents obeying to rules that modify properties in space

Change index (global):

A change index is an indicator derived from multitemporal measurements. It expresses the amount of change within a period of time. It can describe the change behavior of a set of features (global) or of individual features. It can result from a difference, a ratio, ...

Contagious expansion diffusion:

The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion). A **contagious expansion diffusion** is spatially dependant as the phenomenon spreads continuously from a location to contiguous neighbours, in a manner of a contagious disease

Continuous spatial distribution:

Properties of a phenomenon can be distributed throughout space in a continuous or discontinuous manner. A **continuous spatial distribution** is characterised with gentle changes in property values throughout space. This is due to a continuous spatial dependancy throughout space. Thus the spatial distribution corresponds to a continuous surface rather than a set of spatial objects (see Discontinuous spatial distribution)

Cost distance:

When modelling access or movement in space, the usual euclidian distance -called also plane or horizontal distance- is not optimal to express the access to a destination. A most efficient concept is the **cost-distance** that express the amount of resources (financial, time, energy, ...) necessary to move throughout space (see Friction, Barriers)

Densification:

The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The densification expresses an increase in frequency of concerned items or individuals within a specific location or spatial unit of observation. In the process of diffusion of innovation, newly settled locations show an increase in the number of people adopting this innovation, this is known as densification (see Spatial diffusion)

Discontinuous spatial distribution:

Properties of a phenomenon can be distributed throughout space in a continuous or discontinuous manner. A **discontinuous (discrete) spatial distribution** is characterised by strong changes in property values throughout space. Places with value changes delineate spatial feature boundaries (point, linear or areal) (see Continuous spatial distribution)

Extended neighbourhood:

The neighbourhood is an area surrounding a spatial feature or a spatial unit of observation (object or cell). An **extended** neighbourhood generally corresponds to a set of spatial features (objects or cells) that are contiguous to the central feature in a wider sense: they at least share one vertex with it (see spatial neighbourhood and limited neighbourhood)

Forces of movement:

When modelling movement in space, forces and frictions are factors that influence movement. Force can be internal to the moving body, but also external (wind, …). In the computation of **cost-distance**, the influencing forces are modelled as a friction coefficient value attached to each location in space. The friction value is less than 1, but greater than 0 (see friction)

Friction:

When modelling movement in space, impeding properties of space are expressed as frictions. Friction coefficient values can range from 1 (no friction) to a maximum (infinite value) corresponding to an absolute barrier. In the computation of **cost-distance** the assigned friction value is combined with the geometric plane distance. The concept of a friction coefficient can be extended in order to encompass the concept of force. In this case the friction value is less than 1, but greater than 0 (see force of movement)

Game Of Life:

A famous computer simulation illustrating the birth, growth and death of cells distributed throughout a regular gridded space. It demonstrates the use of a **cellular automata model**, using very simple rules, to simulate a spatial diffusion process

Hierarchical diffusion:

The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (**non-spatial dependant diffusion**). In this second situation the diffusion process occurs without spatial continuity, but it is based on the hierarchical importance of centers (places) related with the innovative activity under investigation

High-pass filter:

A technique to **enhance** the local distribution of property values. In the spatial context it makes use of a moving window that defines the neighbourhood. A **gradient** value is computed within the moving window and assigned to the central cell (see low-pass filter)

Isotropic plane surface:

Space can be modelled with different degrees of complexity, such as an isotropic plane surface, an isotropic skewed surface, or an anisotropic skewed surface. The **isotropic plane surface** is the simplest way to model space as a homogeneous plane surface controlled by euclidian geometric properties

Isotropic skewed surface:

Space can be modelled with different degrees of complexity, such as an isotropic plane surface, an isotropic skewed surface or an anisotropic skewed surface. The **isotropic skewed surface** is the second level of complexity to model space as a heterogeneous surface. As each location in space has a specific influence on movements, the euclidian distance concept is replaced with the cost-distance concept

Land Use/Cover Change (LUCC):

The cover or the use of the earth surface is described with a set of category types. Throughout time these cover or use categories can change. The objective of a LUCC analysis is to describe, to understand or to predict such changes in the spatial distribution of landcover or landuse

Limited neighbourhood:

The neighbourhood is an area surrounding a spatial feature or a spatial unit of observation (object or cell). A **limited** neighbourhood generally corresponds to a set of spatial features (objects or cells) that are contiguous to the central feature in a limited sense: they at least share one side with it (see spatial neighbourhood and extended neighbourhood)

Linear regression function:

A regression function that relates a dependant variable Y with one or several independant variables Xi in a linear manner. A first degree polynomial function is a linear function (see Polynomial regression function)

Low-pass filter:

A technique to **smooth** the distribution of property values. In the spatial context it makes use of a moving window that defines the neighbourhood. A central tendancy value is computed within the moving window and assigned to the central cell (see High-pass filter)

Markov chain (analysis):

A technique to estimate the probability of occurrence from any original state to any final state after a specific sequence of n time steps. It makes use of transition matrices

Movements in space:

Actions to move from one location to another through space. The moving feature can be activated by external or internal forces and its path is influenced by local frictions of space

Neighbourhood (spatial):

The neighbourhood is an area surrounding a spatial feature or a spatial unit of observation (object or cell). Based on the spatial dependancy assumption, it stated that the neighbourhood of a feature influences its properties. A neighbourhood is defined by a range and a shape that encompass the feature up to a certain distance

Non-spatially dependent diffusion:

The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (**non-spatial dependant diffusion**). In this second situation the diffusion process occurs without spatial continuity

Path (trajectory):

The trajectory used to move from an origin to a destination in space

Polynomial function:

A non-linear function that relates a dependant variable Y with one or several independant variables Xi. This function contains a combination of polynomes at different degrees or exponents (see Linear regression function)

Prediction stage:

The application of a predictive model is made of 2 stages: calibration and **prediction** stages. Once the model has been calibrated to suit the period of time validated with measurements, it can be used to predict feature properties estimated during a period of time with no existing measurements (generally a future period of time)

Relocation diffusion:

A type of spatial diffusion process where previous locations of presence are replaced by new locations throughout time

Scenarios of development:

A scenario of development describes the possible evolution of properties of a phenomenon, based on specific rules of change and specific characteristics (parameters) at the starting date. A set of different scenarios can be defined in order to simulate possible alternatives of development

Sleuth (model):

The SLEUTH name is the acronym for Slope, Landuse, Exclusion, Urban, Transportation and Hillshading. The Sleuth model is a cellular automata simulation model developed to model the development of urban areas (Clarke et al., 1998)

Spatial diffusion:

The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (spatial dependant diffusion) or independant (non-spatial dependant diffusion)

Spatial dynamics modelling:

Procedure used to summarise the changes in the spatial distribution of phenomenon properties throughout time (see Spatial dynamics)

Spatial expansion:

The impact of the presence of a phenomeon throughout space. In a spatial diffusion process, the phenomenon spatially expands throughout time

Spatial filtering:

Procedures used to isolate spatial components: local and regional variations, trend. Typical filtering techniques make use of a moving window for the smoothing of local variations (low-pass filters) or their enhancement (high-pass filters)

Spatially dependent diffusion:

The process of change in the spatial distribution of thematic properties throughout time can be approached from different viewpoints: a contiguous or dispersed spread of presence, a densification, … The diffusion process can be spatially dependant (**spatial dependant diffusion**) or independant (non-spatial dependant diffusion). In this second situation the diffusion process occurs with spatial continuity

Temporal variability:

Some properties of space or spatial features can change throughout time, as others remain constant. Their temporal variability is then different

Trend surface (analysis):

A regression function modelling the property values (Z) based on their location (X,Y) in space: Z = f(X,Y)

1.8. Bibliography

- **Abler R., Adams J., Gould P.**, 1972. *Spatial Organization, The Geographer's View of the World*. USA: Prentice Hall.
- **Al-Ghamdi, K.**, 2008. *Land Use Change Modeling, A Cellular Automata Simulation Approach, Concepts and Applications, Case Study of Bulle, Switzerland*. (PhD Thesis Nr. 1617). University of Fribourg, Switzerland.
- **Batty, M. and Y. Xie**, 1994. From cells to cities. *Environment and Planning*, B 21, 41-48.
- **Caloz R., Collet C.**, 2011. *Analyse spatiale de l'information géographique*. Lausanne, Switzerland: PPUR.
- **Cecchini A. e Rinaldi E.**, 1999. The Multi-Cellular Automaton: a tool to build more sophisticated models. A theoretical foundation and a practical implementation. *In:* **Rizzi P.**, ed. *Computer in Urban Planning and Urban Management 6th International Conference*. Milano, FrancoAngeli.
- **Clarke, K. C., S. Hoppen and L.J. Gaydos**, 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning*, 247-261.
- **Clarke, K.C. and Gaydos, L.J.**, 1998. Loose-coupling a cellular automata model and GIS:long-term urban growth prediction for San Francisco and Washington/Baltimore. *Int. J. Geographical Information Science*, 12 n°7, 699-714.
- **Cliff, A. D., P. Hagget, J. K. Ord and G. Versey**, 1981. *Spatial Diffusion : An Historical Geography of Epidemics in an Island Community*. Cambridge University Press.
- **Conway, J.**. *The Game of Life* [online]. Available from: http://www.bitstorm.org/gameoflife.
- **Davis, J.C.**, 1986. *Statistics and data analysis in geology*. New York: John Wiley &Sons.
- **Eastman, R.**, 2008. *IDRISI - Taiga GIS and Image processing software, Reference Manual*. Worcester, Clark Labs, Clark University, USA: Worcester.
- **Frank, L.**, 2002. A spatial model of the diffusion of mobile communications within the European Union. *In: European Regional Science Association/ERSA conference papers*.
- **Gould, P.**, 1969. *Spatial Diffusion*. USA: Association of American Geographers, Commission on College Geography Resource Paper No.4.
- **Hägerstrand, T.**, 1967. *Innovation Diffusion as a Spatial Process*. Lund, Sweden: Gleerups.
- **Li, X. and A. G. Yeh**, 2000. Modeling sustainable urban development by the integration of constrained cellular automata and GIS. *Int. J. Geographical Information Science*, 14(2), 131-152.
- **Saint_Julien, T.**, 2001. *Processus de diffusion spatiale et modélisation du changement. In Chapitre 5, Modèles en analyse spatiale*. Paris, France: Hermès Science Publication (dir. Sanders L.). [pp. 157-186.]
- **Singh, A. K.**, 2003. *Modelling Land Use Land Cover Changes using Cellular Automata in a Geo-Spatial Environment*. (Master Thesis of Science in Geoinformatics). ITC, Enschede, The Netherland.
- **Takeyama, M. and H. Couclelis**, 1997. Map dynamics: integrating cellular automata and GIS through Geo-Algebra. *Int. J. Geographical Information Science*, 11(1), 73-91.
- **Tim Tyler**. *Reversible Cellular Automata, neighborhoods, billiard ball machines, Java applets* [online]. Available from: http://cell-auto.com.
- **Treuil, J-P, Mullon, Ch, Perrier, E., Piron, M.**, 2001. *Simulations multi-agents de dynamiques spatialisées. In Chapitre 7, Modèles en analyse spatiale*. Paris, France: Hermès Science Publication (dir. Sanders L.). [pp. 219-249.]
- **Unwin, D.**, 1975. *An Introduction to Trend Surface Analysis*. London, GB: Catmog 5.